

هوش تجاری

Business Intelligence

دکتر مهدی اسماعیلی

کتاب منابع هوش تجاری:

- ۱- هوش تجاری؛ مترجم دکتر محمدرحیم اسفندانی و دکتر روح اله نوری - دانشگاه تهران
- ۲- هوش تجاری؛ تألیف: دکتر تارخ ومهاجری - دانشگاه خواجه نصیرالدین طوسی
- ۳- هوش تجاری موفق؛ مترجم دکتر علاقه بند Rapid Miner - Weka - Shair point

مقدمه:

هوش تجاری به سه بخش تقسیم می‌شود:

- ۱- کنترل و مدیریت پروژه
 - ۲- انبارسازی داده‌ها (Data Warehousing)
 - ۳- داده کاوی (Data Mining)
- چگونگی اجرای یک پروژه BI در یک سازمان، تهیه گزارشات مدیریتی
مباحث مربوط به جمع‌آوری و انبارش داده‌ها
روش‌های تحلیل هوشمند داده‌ها

تعاریف متفاوت از هوش تجاری

- هرگونه کار و عملیاتی که در یک سازمان انجام شود و هرگونه تفکر مدیریتی (استخدام ، تغییر پست افراد ، جابجایی اتاق‌های افراد) و هرگونه ابزار مدیریتی که در سازمان‌ها به کار برده می‌شود و موجب رشد قابل توجه کارایی سازمان شود ، هوش تجاری گویند.
- هوش تجاری مدیریت درست و مؤثر اطلاعات برای هدفی تجاری است و اساساً به ارتباط بین کاربر و سیستم‌های فناوری اطلاعات مربوط می‌شود.
- هوش تجاری عبارت است از شرح نیازهای سازمان برای تجزیه و تحلیل به همراه گزارش‌دهی.
- هوش تجاری رویکردی راهبردی است برای هدف‌گذاری به صورت سیستماتیک، ردیابی، مخابره و نهایتاً تبدیل و ترجمه سیگنال‌ها و علائم ضعیف کسب‌وکار به یکسری اطلاعات کاربردی که اساس تصمیم‌گیری‌های راهبردی شوند.
- هوش تجاری به عنوان تغییرشکل‌دهنده‌ای آگاه و روشمند است که داده‌های مختلف را از منابع مختلف دریافت و به شکل‌های جدید تغییر می‌دهد تا اطلاعاتی نتیجه‌گرا و تجاری را فراهم کند.
- هوش تجاری مبحث بسیار وسیعی است که کارکردها (مانند گزارش‌دهی و تجزیه و تحلیل) و تکنولوژی‌های متفاوتی (مانند انبار داده، OLAP ، پورتال) را در بر می‌گیرد.

- هوش تجاری شامل تمام ابزارهای اطلاعاتی و تکنیکی جهت تحلیل دانش و اطلاعات در دسترس یک شرکت است و به طور کلی عبارت است از دسترسی، تجزیه و تحلیل و فراهم آوردن داده‌های تجاری برای کاربران در یک بنگاه تجاری.
- هوش تجاری ابزاری است که توسط بنگاه‌ها برای جمع‌آوری، مدیریت، تجزیه و تحلیل داده‌ها و اطلاعات اساسی و ساخت‌یافته و یا غیر آن استفاده میشود و این کار با به کارگیری قابلیت‌های فناوری اطلاعات صورت می‌پذیرد.
- هوش تجاری یک معماری است و مجموعه‌ای از برنامه‌های کاربردی و پایگاه داده‌های تصمیم‌یار و عملیاتی یکپارچه است که دسترسی به داده‌های تجاری را آسان می‌کند.

اصطلاحات ویژه در مباحث هوش تجاری

- Transaction Processing System (TPS) : سیستم پردازش تراکنش
- Office Automation Systems (OAS) : سیستم‌های اتوماسیون اداری
- TPS و OAS سیستم‌هایی هستند که مدیران سطح پایین با آنها کار می‌کنند.
- Individual Information System (IIS) : سیستم اطلاعات شخصی
- Organizational Information Systems (OIS) : سیستم های اطلاعات سازمانی
- Information System (IS) : سیستم اطلاعات
- Management Information System (MIS) : سیستم‌های اطلاعات مدیریتی (MIS و DDS) این سیستم‌ها، سیستم‌های هوش تجاری نیستند بلکه بخشی از آن هستند.
- Decision Support System (DSS) : سیستم‌های پشتیبان تصمیم‌گیری
- سیستم‌هایی که کمک می‌کنند تا تصمیماتی را گرفته شوند. مانند نرم‌افزارهایی که گزارش‌های خاصی را می‌گیرند ، مثل نرم‌افزار Excel در این نرم‌افزار می‌توان گزارش‌های متنوعی را تهیه کرد، با این نرم‌افزارها می‌توان نمودارهایی رسم کرد که بتوان با آنها به پشتیبانی از تصمیم پرداخت.
- Strategic Information Systems (SIS) : در این سطح تقسیمات راهبردی را خواهیم داشت در این دو سطح بالا با Information درگیر هستیم.
- Enterprise Information System (EIS) : معمولاً این سیستم‌ها گزارش‌های خیلی حرفه‌ای می‌گیرند که معمولاً این گزارش‌های حرفه‌ای را با Query های SQL نمی‌توان دریافت کرد و ساخت. معمولاً این گزارشات با جدول‌هایی که دو بعدی هستند جواب‌گو نیستند.

رده‌های مختلف مدیران و اهمیت تصمیم‌گیری

هر مدیر برای مدیریت درست و کارآمد علاوه بر همه قابلیت‌های مختلف مدیریتی ، مهم‌ترین چیزی که برای مدیریت صحیح و کارآمد نیاز دارد « داده‌ها » است.

در هرم مدیریت:

- در رده‌ی TPS و OAS ، اغلب مدیرانی هستند که کارهای دفتری انجام می‌دهند. مثل کسانی که سند حسابداری تنظیم می‌کنند یا مشخصات افراد را وارد رایانه می‌کنند.

- مدیران سطح میانی؛ گزارشات مدیریتی را مورد بررسی قرار داده و تصمیمات تاکتیکی می‌گیرند.
- مدیران ارشد؛ تصمیمات راهبردی را اتخاذ می‌کنند و چشم‌انداز سازمان را تعیین می‌کنند.

در این هرم BI در همه جا کاربرد دارد ، اما بیشتر از سایر بخش‌ها BI مورد نیاز مدیران اجرایی ارشد است به جهت این‌که این مدیران برای اتخاذ تصمیمات استراتژیک نیاز به داده‌های صحیح دارند و BI این اطلاعات را برای آنها فراهم می‌کند.

توجه: هدف هوش تجاری این است که داده‌های درست در زمان درست به شخص درست تحویل داده شود.

برنامه‌های کاربردی تصمیم‌محور در هوش تجاری بسیاری از فعالیت‌ها را آسان می‌کنند:

- تجزیه و تحلیل چندبعدی، مثل: پردازش تحلیلی آنی (آنلاین) (OLAP)
- تجزیه و تحلیل کلیک محور
- داده‌کاوی (کاوش متن ، محتوا و صوت)
- پیش‌بینی
- تجزیه و تحلیل کسب و کار
- فراهم نمودن کارت امتیازی متوازن
- مصور کردن
- پرس و جو ، تهیه گزارش و رسم نمودار (شامل هشدارهای عامل محور و بر اساس زمان)
- تجزیه و تحلیل مکانی
- مدیریت دانش
- پیاده‌سازی پورتال
- دسترسی داشبورد دیجیتال
- دیگر فعالیت‌های متقابل کارکردی

مثال هایی در رابطه با پایگاه های داده ی هوش تجاری تصمیم محور

- شرکت انبار حجم گسترده ای از داده ها
- انبار داده ها (کاربرد ی و بخشی)
- انبار اکتشافی (آماری)
- پایگاه های داده کاوی
- انبارهای وب (برای اطلاعات کلیک محور)
- فروشگاه داده های عملیاتی (ODSs)
- انبارک های عملیاتی (Oper Marts)
- دیگر فعالیت های متقابل پایگاه های داده ی تصمیم محور

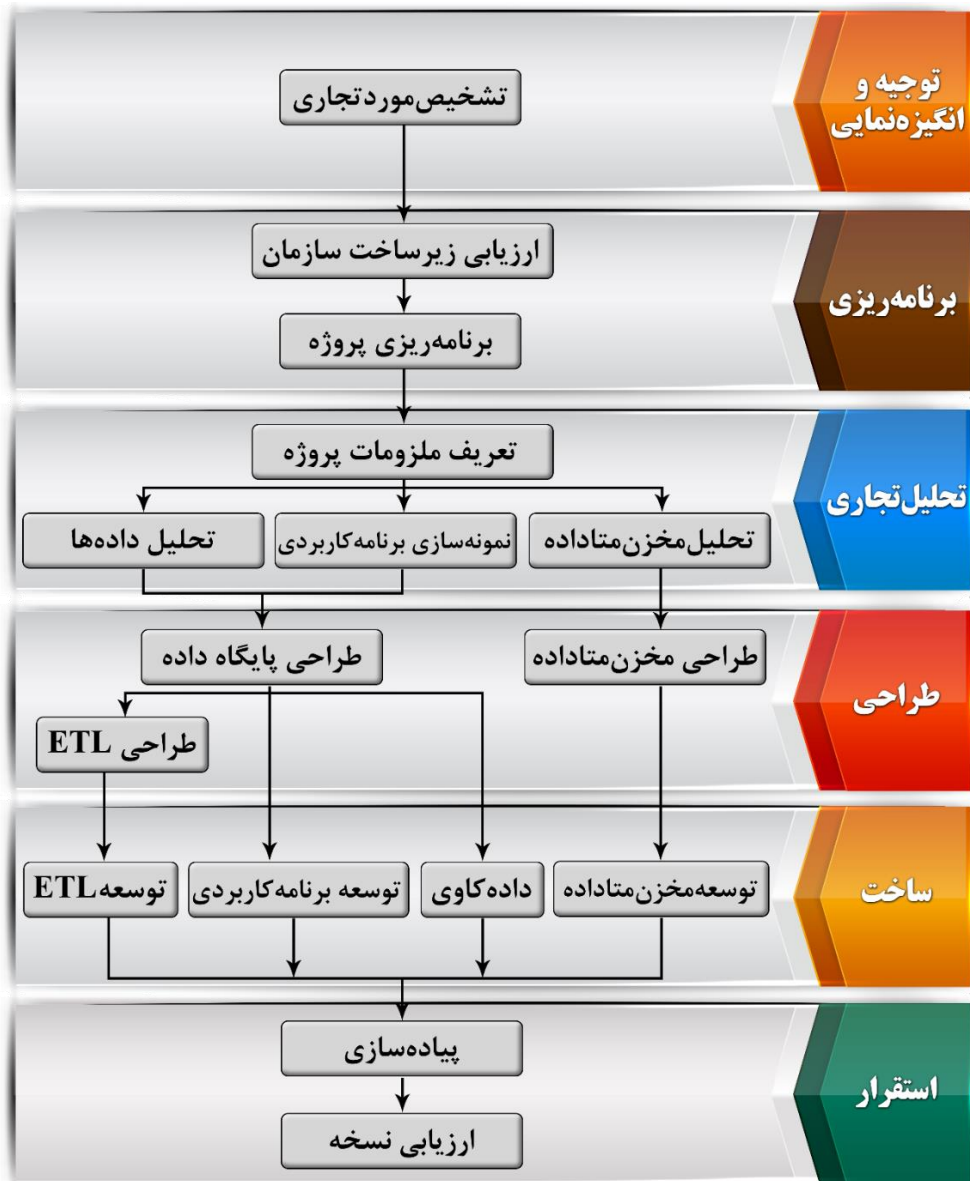


نقشه راه هوش تجاری (گام های هوش تجاری):

در هر پروژه BI فازها و گام هایی برای راه اندازی پروژه در یک سازمان یا شرکت متصور است. مراحل و گام های لازم برای راه اندازی پروژه BI شامل ۶ مرحله و ۱۶ گام (Step) است. که عبارتند از:

مرحله اول (توجیه و انگیزه نمایی): هر پروژه ای که می خواهد آغاز شود ابتدا باید چرایی آن پروژه بررسی و تعریف شود. در پروژه حتماً باید توجیه و انگیزه ای وجود داشته باشد؛ زیرا مجموع این توجیه ها همان هدفی است که ارزشیابی می شود ، وقتی توجیهی وجود ندارد پارامتری هم برای ارزشیابی وجود نخواهد داشت. در تشخیص مورد تجاری باید از افراد خاصی کمک گرفت و از همه مهم تر خروجی این مرحله است که باید ارزشیابی شود.

مرحله دوم (برنامه ریزی): در این فاز زیرساخت های فنی و غیر فنی بررسی می شوند. یعنی این که چه امکاناتی وجود دارد و کدام یک از آنها مورد استفاده هستند. (مثلاً نیاز به حافظه های با حجم بالا که نیاز به یک زیرساختار فنی دارد، که ممکن است محل انجام این پروژه این امکانات را داشته یا نداشته باشد).



توجه: در هر گام باید به ۴ موضوع رسیدگی شود:

- ۱- در هر گام چه کارهایی انجام می‌شود؟ یعنی Activity یا فعالیت.
- ۲- افراد درگیر در هر گام چه کسانی هستند.
- ۳- خروجی این گام چیست؟ یعنی خروجی خاص هر گام شامل چه چیزهایی هست.
- ۴- ریسک‌های پروژه

مرحله سوم (تحلیل تجاری): در این گام باید ملزومات مورد نیاز پروژه مشخص شود و این بعد از ارزشیابی زیرساخت‌های سازمان قرار می‌گیرد که با تحلیل در زیرساخت‌ها در این گام مشخص می‌شود به چه چیزهایی نیاز است و چه چیزهایی باید تهیه نمود و این زیرساخت‌ها در کجای پروژه نیاز است.

پیشگونه‌سازی (Prototyping): ساخت یک نمونه ابتدایی از پروژه نهایی است که قبل از انجام پروژه انجام می‌گیرد.

حسن این گام قرارداد است؛ یعنی برای اینکه بتوان درمورد قرارداد به توافق رسید باید از یک Prototype استفاده کرد. برای تعریف User Interface باید پیشگونه سازی کرد.

در پیشگونه‌سازی باید ابزارها را هم به خوبی شناخت و فقط تعریف آن مورد نیاز نیست.

تحلیل داده‌ها: یکی از مراحل انجام انبارسازی داده‌هاست.

تحلیل مخزن متاداده‌ها: در واقع Data about Data است. یعنی تعریفی برای داده‌هاست و اطلاع‌هایی که بیشتر در حوزه داده‌ها وجود دارد. وقتی پروژه‌ای شروع می‌شود به Data نیاز است و داده‌ها در یک محل متمرکز نیست.

متادیتا اطلاعات و داده‌هایی است که درباره‌ی داده‌های موجود توضیح می‌دهد ، اما الزاماً داده‌های مورد نظر را تعریف نمی‌کند.

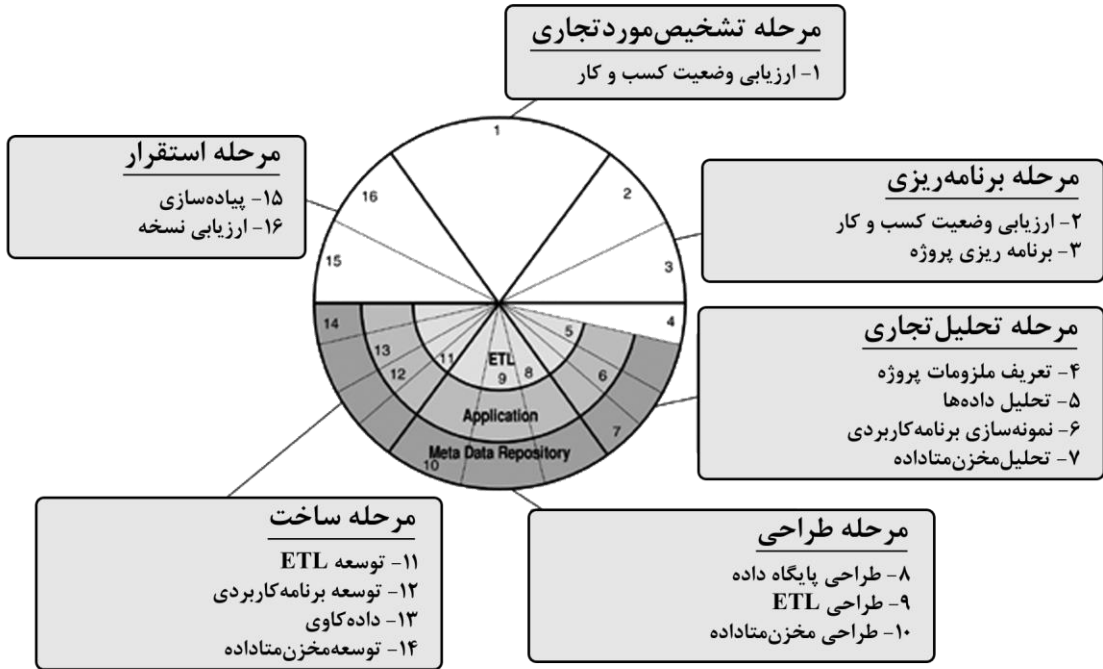
یکی از مزایای متادیتا جستجوی سریع است و سرعت بخشیدن به Query ها یا پرسش‌های کاربران و اطلاعات در شناسنامه داده‌ها، شاخص‌ها که داده‌های مورد نظر نیستند. مزیت دیگر داشتن متادیتا این است که امنیت و زنده بودن پروژه را تضمین می‌کند و شکل ذخیره سازی متادیتا نیز اهمیت زیادی دارد. (سرچ شود Meta Data Repository)

طراحی ETL : (Extract/Transform/Loading) یکی از کارهایی که در هوش تجاری انجام می‌شود، یکپارچه سازی داده‌هاست؛ یعنی داده‌ها کنار هم قرار بگیرند زیرا داده‌ها ممکن است ایراداتی مانند تکرار ، missing ، هم واحد نبودن را داشته باشند پس فرآیند بزرگ ETL این مشکلات را رفع می‌کند. **Extract:** یعنی داده‌ها از کجا استخراج شوند. (وزن مشخص)

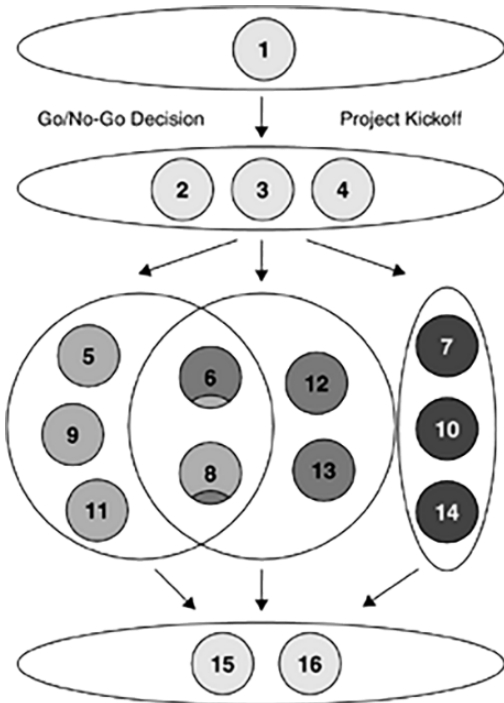
Transform: یعنی داده‌ها از شکلی به شکل دیگر تبدیل شوند.

Loading: یعنی بارگذاری روی رسانه‌ای که (Data Warehousing) روی آن بارگذاری می‌شود. (بخش اول می‌تواند به‌عنوان پروژه باشد). (سرچ BI tools)

بخش‌های دارای قابلیت توسعه موازی (برای گام‌های ۵ تا ۱۴)



مراحل انجام شده در بخش‌های توسعه موازی



○ مراحل قبل و بعد از انجام پروژه به بخش‌های توسعه موازی تقسیم می‌شوند. (گام‌های ۱، ۲، ۳، ۴، ۱۵ و ۱۶)

● مراحل در مسیر انجام بخش ETL. (گام‌های ۵، ۸، ۹، ۱۱ و مشارکت جزئی در مرحله ۶)

● مراحل در مسیر انجام بخش برنامه کاربردی. (گام‌های ۶، ۱۲، ۱۳ و مشارکت جزئی در مرحله ۸)

● مراحل در مسیر انجام بخش مخزن متاداده. (گام‌های ۷، ۱۰ و ۱۴)

افراد و تیم پروژه‌های هوش تجاری

به‌طور کلی در یک پروژه هوش تجاری در ۶ گام آن ۱۱ نقش اصلی و ۱۶ نقش کمکی وجود دارد:
به دو دسته می‌شوند: ۱- تیم اصلی ، ۲- تیم فرعی

۱- تیم اصلی: که شامل ۱۱ نقش است و به آن Core Team یا هسته می‌گویند.

تیم اصلی، از افرادی تشکیل می‌شوند که این افراد در طول پروژه ، در هسته و در تمام گام‌های پروژه حضور دارند. البته نقش آنها در بعضی گام‌ها پررنگ تر و در بعضی کم رنگ‌تر است.
معمولاً در سطوح بالا مدیران اجرایی و ارشد بیشتر کار فکری می‌کنند و با گزارشات مدیریتی درگیرند و با گزارشات حجیم و داده‌های آماری و ریز درگیر نیستند ؛ یعنی باید بتوانند گزارشات را باهم مقایسه کنند. بدون اینکه خیلی در آنها ریز شوند این افراد در این سطح نخبه هستند اما این نشان دهنده کم اهمیت بودن مدیران عملیاتی نیست و پایین بودن آنها به دلیل افراد بیشتر در این سطح است.

نقش و مسئولیت تیم هسته

نقش	مسئولیت‌های اصلی
مسئول توسعه برنامه	طراحی و نظارت بر توسعه‌ی برنامه (مثل: گزارشات و تحقیقات)
معمار زیرساخت BI	ایجاد و حفظ زیرساخت های فنی BI
نماینده کسب و کار (نماینده تجاری)	شرکت در جلسات مدل سازی، ارائه تعاریف داده‌ها، نوشتن موارد آزمون، تصمیم گیری کسب و کار، حل و فصل اختلافات بین واحدهای کسب و کار، و بهبود کیفیت داده‌ها تحت کنترل واحد کسب و کار
مدیر اطلاعات	انجام تجزیه و تحلیل داده‌های متقابل سازمانی، ایجاد مدل‌های منطقی داده‌ها پروژه‌های خاص، و ادغام مدل منطقی داده‌ها در یک مدل منطقی تجاری داده‌ها
کارشناس داده‌کاوی	انتخاب و اجرای ابزار داده‌کاوی (باید زمینه آماری داشته باشد)
تحلیلگر کیفیت داده‌ها	بررسی کیفیت منبع داده و فراهم نمودن مشخصات پاکسازی داده برای فرایند ETL
مدیر پایگاه داده	طراحی، بارگذاری، نظارت، و تنظیم پایگاه داده هدف BI
مسئول توسعه ETL	طراحی و نظارت بر فرآیند ETL
مدیر متاداده	ایجاد و یا صدور مجوز (خرید)، افزودن، بارگیری، و حفظ مخزن داده‌های متا
مدیر پروژه	تعریف، برنامه ریزی، هماهنگی، کنترل و بررسی تمام فعالیت‌های پروژه ، ردیابی و گزارش پیشرفت ، حل و فصل مسائل فنی و کسب و کار ، مشاوره تیم مذاکره با فروشندگان ، نمایندگی کسب و کار ، حامی کسب و کار؛ مسئولیت کلی برای پروژه
کارشناس موضوع اصلی	ارائه اطلاعات کسب و کار و نیازمندی‌ها در مورد داده‌ها و فرآیندها

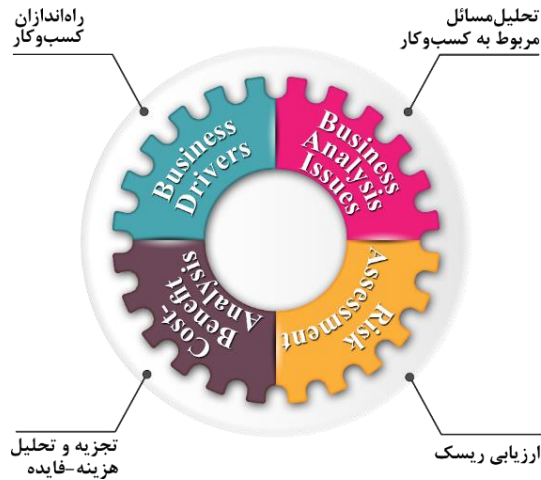
۲- تیم کمکی: تیمی است که لزومی به استفاده از آنها در تمام مراحل پروژه نیست و فقط برای تضمین اطمینان ، از آنها استفاده می‌شود.

نقش‌ها و مسئولیت‌های گسترش یافته تیم (تیم فرعی)

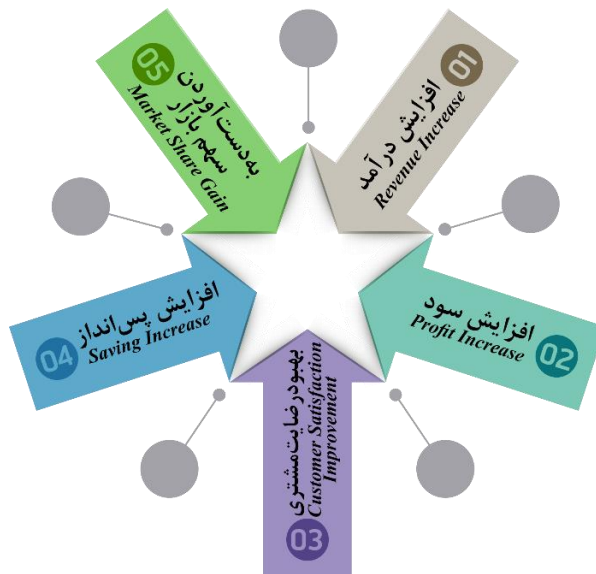
نقش	مسئولیت‌های اصلی
توسعه دهنده(گان) برنامه	برنامه نویسی برنامه گزارش، نوشتن اسکریپت پرس و جو ، و توسعه دسترسی و تجزیه و تحلیل برنامه‌های کاربردی
پشتیبان BI (همکاران کمکی)	مشاوره و آموزش کارکنان کسب و کار
حمایت کننده تجاری	دفاع از ایجاد BI و از بین بردن موانع مربوط به کسب و کار برای تیم پروژه BI
توسعه دهنده(گان) ETL	برنامه نویسی برنامه‌های ETL و / یا آماده نمودن ابزارهای ETL
حسابرس IT و یا تحلیلگر QA	تعیین خطرات و مواجهه پروژه BI به علت عدم کنترل نیروهای داخلی یا خارجی
توسعه دهنده(گان) مخزن متاداده	برنامه نویسی برنامه‌های مهاجرت مخزن داده‌های متا برای بارگذاری پایگاه داده مخزن متاداده. ارائه گزارش داده‌های متا و کمک رسانی آنلاین
کارکنان خدمات شبکه	حفظ و نگهداری از محیط و فضای شبکه
کارکنان عملیاتی	اجرای فرآیندهای دسته ای برای چرخه ETL ، دسترسی و تجزیه و تحلیل نرم‌افزار، و مخزن داده‌های متا
سرپرست امنیتی	تضمین این که نیازهای امنیتی تعریف شده است و ویژگی های امنیتی در سراسر همه ابزار و پایگاه‌های داده مورد آزمایش قرار گرفته است
سهام‌داران (دیگر نمایندگان کسب و کار و یا مدیران IT)	مدیریت محدوده‌ی مسئولیت‌ها در پروژه BI ، مانند بررسی و تصویب استانداردهای متقابل سازمانی و قوانین و اساسنامه کسب و کار که تیم پروژه BI استفاده می‌کند و یا آن را ایجاد می‌کند.
معمار استراتژیک	مدیریت زیرساخت های فنی کلی برای سازمان، از جمله زیرساخت های فنی BI
کارکنان خدمات فنی	حفظ زیرساخت های سخت افزاری و سیستم عامل
آزمایش کنندگان	آزمایش کدهای برنامه‌نویسی ایجاد شده توسط توسعه دهندگان ETL ، کاربرد، و بخش‌های مخزن متاداده
مدیران ابزار	نصب و نگهداری از ابزارهای توسعه و دسترسی و ابزار تجزیه و تحلیل
توسعه دهنده(گان) وب	طراحی وب سایت و ایجاد صفحات وب برای نمایش گزارش و نمایش داده‌ها در اینترنت، اکسترانت، و یا اینترنت
کارشناس ارشد وب	راه‌اندازی وب سرور و وب سایت‌های امنیتی

گام اول: توجیه و انگیزه نمایی اجزای کسب‌وکار

در این مرحله اهداف استراتژیک کسب و کار سازمان واضح و روشن می شود



دسته بندی اهداف (معمولاً پروژه های IT یکی از پنج هدف زیر را دنبال می کنند)



۱- افزایش درآمد

- شناسایی بازارهای جدید و بخش های ویژه بازار
- انگیزه ی فروش مؤثرتر
- شناسایی سریع تر فرصت ها
- زمان انتقال سریع تر به بازار

۲- افزایش سود

- هدف گذاری بهتر ایمیل های تبلیغاتی
- واکنش سریع به کاهش بازارها
- شناسایی محصولات یا خطوط تولید آنها
- شناسایی ناکارآمدی داخلی
- مدیریت کارآمدتر کالا

۳- بهبود رضایت مشتری

- درک بهتر از ترجیحات مشتری
- بهبود محصول مطابق نظر مشتری
- فروش فوق العاده برای مشتریان
- افزایش کسب و کار تکراری (مشابه)
- راه حل سریع تر با بررسی شکایت مشتری

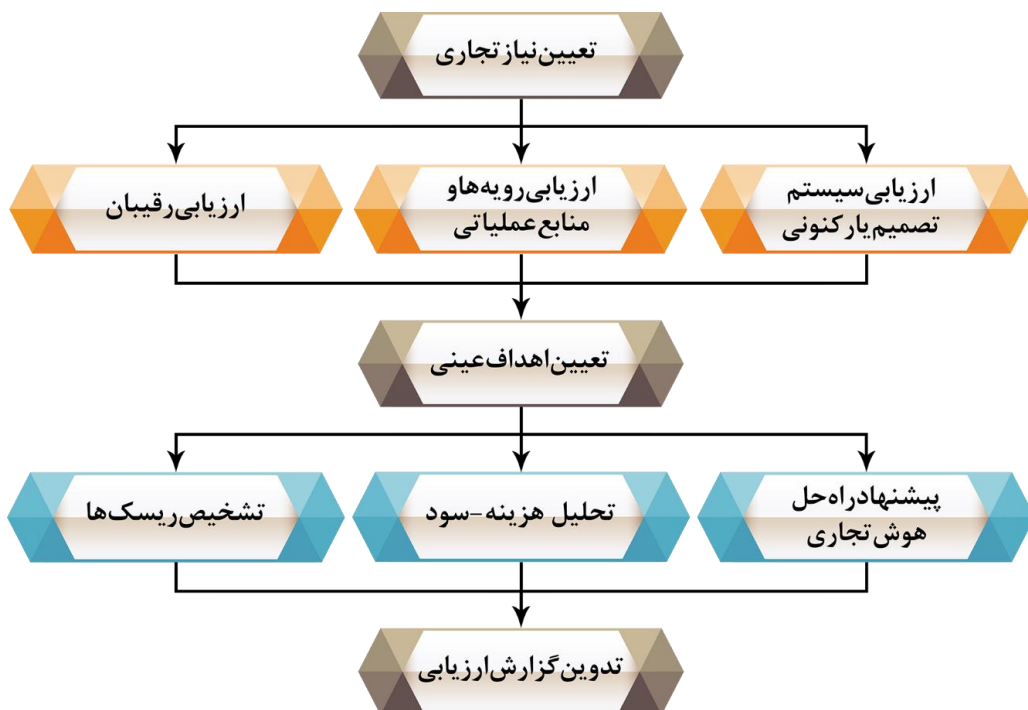
۴- افزایش پس انداز

- کاهش کالای تلف شده یا از تاریخ گذشته
- کاهش درخواست گزارش های سفارشی

۵- به دست آوردن سهم بازار

- افزایش تعداد مشتریانی که به واسطه ی ایراد در نحوه ی رقابت از دست می روند
- حفظ بسیار بالاتر میزان افزایش مشتری در مقایسه با رقابت در سال های گذشته

موارد ارزیابی فعالیت های کسب و کار: فعالیت هایی که در گام اول باید انجام شوند



تعیین نیاز تجاری (تعیین اهداف پروژه): برای شروع یک پروژه هوش تجاری باید دلیل کار کاملاً واضح و روشن باشد و حتی المقدور توضیحات صاحبان سازمان صورت جلسه شود پس باید دانست این پروژه چه دردی را می خواهد دوا کند.

ارزیابی سیستم تصمیم‌یار کنونی: وقتی یک پروژه هوش تجاری انجام می‌شود الزاماً صفر کیلومتر نیست یعنی قبلاً در این سازمان خیلی کارها انجام شده و سیستم‌های اتوماسیون در آن سازمان وجود دارد.

تعیین اهداف عینی: اهداف باید کاملاً روشن شود. یکی از اهداف افزایش سود است. (Benefit)

جلب رضایت مشتری: یکی از اهداف پروژه‌های هوش تجاری است. جلب رضایت مشتری‌ها افزایش سود را به دنبال دارد، عدم رضایت مشتری و عدم شفافیت در آمار و ارقام و گزارش‌ها عدم افزایش سود را به دنبال دارد.

ارزیابی ریسک

ریسک: تصمیم‌گیری برای عدم قطعیت آینده‌ی پروژه را ریسک می‌نامند.

مثلاً در ابتدای پروژه با توجه به سابقه‌ی یک فرد خاص در پروژه حدس زده می‌شود ممکن است این فرد پس از مدت مشخصی کار را رها کند، بنابراین برای این اتفاق تصمیم‌گیری شود که اگر این فرد پروژه را ترک کرد چه کسی جایگزین او در پروژه شود، این کار تحلیل ریسک است. ریسک‌ها به دسته‌های مختلفی تقسیم می‌شوند برخی ریسک‌ها احتمال وقوعشان کم است ولی تأثیرشان خیلی زیاد است. اما برخی ریسک‌ها احتمال وقوعشان بالاست اما تأثیرگذار نیستند.

ریسک‌ها دارای دو پارامتر اصلی هستند:

۱- درصد وقوع ریسک (درصد بالا - پائین) ۲- میزان تأثیر ریسک (میزان تأثیر بالا - پائین) با ترکیب این دو پارامتر، چهار نوع ریسک وجود دارد.

ارزیابی ریسک

- فن‌آوری مورد استفاده برای اجرای پروژه
- پیچیدگی قابلیت‌ها و فرآیندها برای به اجرا درآوردن پروژه
- ادغام اجزای مختلف و داده‌ها
- سازمان و حمایت مالی و معنوی آن
- مهارت‌ها، نگرش‌ها، و سطح تعهد کارکنان تیم این پروژه
- سرمایه‌گذاری مالی از نظر بازدهی سرمایه

ماتریس پایه ارزیابی ریسک

سبز = کم خطر، حرکت رو به جلو همراه با پروژه

زرد = خطر متوسط (احتیاط)، ادامه پروژه به آرامی و با احتیاط

قرمز = خطر بالا (توقف)، ارزیابی مجدد قبل از هرگونه اقدام

سطح ریسک			
خطر بالا	خطر متوسط	کم خطر	متغیر
فناوری‌های جدید، تجربه کم	تجربه حداقلی در مورد فناوری	آشنایی کامل با فناوری	فناوری
مأموریت بحرانی ، نیازمند مهندسی مجدد گسترده	متوسط، تأثیر موردی در گردش کار	ساده، تأثیر حداقل در گردش کار	پیچیدگی
نیازمند ادغام گسترده	نیازمند ادغام محدود	مستقل، بدون ادغام	ادغام
پشتیبانی داخلی محدود	تا حد زیادی حمایتی	پشتیبانی داخلی ثابت	سازمان
بدون تجربه‌ی کسب و کار، تنها تکنولوژی محور ، استعداد محدود ، با نگرش منفی	کم تجربه در کسب و کار، کسب و کار راه اندازی می‌شود ، با استعداد ، با نگرش عادلانه	با تجربه در کسب و کار، کسب و کار راه اندازی می‌شود ، با استعداد ، با نگرش باز	تیم پروژه
امکان بازگشت سرمایه طی چندین سال	بازگشت احتمالی سرمایه در بازه زمانی متوسط	بازگشت احتمالی سرمایه در مدت زمان بسیار کوتاه	سرمایه‌گذاری
ناقص، نمی‌تواند ارزیابی شود	کامل و تنظیم نشده	کامل و تنظیم شده	مدل داده‌های فیزیکی: کامل
ناقص، نمی‌تواند ارزیابی شود	توسعه نوع محدودی از محصول	توسعه کامل برای انواع محصول جدید	مدل داده‌های فیزیکی: توسعه انواع محصول جدید
ناقص، نمی‌تواند ارزیابی شود	عملکرد و یا نگرانی‌های زمان بندی	پشتیبانی طراحی قابل قبول برای سیستم‌های منبع	مدل داده‌های فیزیکی: تغذیه سیستم منبع
پشتیبانی ضعیف برای رابط خارجی و داخلی	پشتیبانی محدود برای رابط خارجی و داخلی	پشتیبانی از رابط‌های خارجی و داخلی	رابط (خارجی و داخلی)
در حال حاضر نمی‌توان آن را ارزیابی کرد	می‌تواند اضافه شود، اما نیاز به بازسازی قابل توجهی برای مکعب دارد	آسان برای اضافه کردن	ابعاد تجزیه و تحلیل و اقدامات: اضافه کردن خطوط تولید جدید
مکعب‌های پیشنهادی و مجموعه‌ای از ابعاد کافی	مکعب‌های پیشنهادی و مجموعه‌ای از ابعاد ارائه با حداقل کفایت	مکعب‌های پیشنهادی و مجموعه‌ای از ابعاد برای حمایت کافی از تحلیلگران کسب و کار	ابعاد تجزیه و تحلیل و اقدامات: اضافه کردن ابزارهای جدید برای تجزیه و تحلیل داده‌ها
بدون حمایت متا داده	پشتیبانی متا داده محدود	به طور کامل توسعه یافته	استفاده از مخزن داده‌های متا
بارگذاری فرایندها انجام نشده و نمی‌تواند ارزیابی شود	بارگذاری ضعیف فرایندها یا اسناد	بارگذاری کامل فرایندهای ایجاد شده	بارگیری پایگاه‌های داده هدف BI
طراحی پایگاه داده‌های فیزیکی ناقص، نمی‌تواند ارزیابی شود	مشکلات جزئی با طراحی پایگاه داده‌های فیزیکی	طراحی پایگاه داده‌های فیزیکی مؤثر و کارآمد	مشکلات پایگاه داده‌های فیزیکی

در حال حاضر نمی توان آن را ارزیابی کرد	برخی از مسائل مربوط به عملکرد	مطابق با اعلام عملکرد مورد نیاز	مسائل مربوط به عملکرد
هیچ پشتیبانی از فرایندها نمی شود	پشتیبانی محدود مستندات	پشتیبانی فرایندها به خوبی مستند و تثبیت شده	مسائل مربوط به مدیریت سیستم: تعمیر و نگهداری
درمورد پشتیبان گیری و بازیابی فرایندها فکر نشده	پشتیبان گیری و بازیابی فرایندهای توسعه یافته اما نصب نشده	پشتیبان گیری و بازیابی فرایندهای توسعه یافته و نصب شده	مسائل پشتیبانی
طراحی امنیت ناقص، نمی تواند ارزیابی شود	برای نگهداری دشوار است	برآورده ساختن نیازهای نرم افزاری که به سادگی قابل نگهداری هستند	اجرای امنیت

نتایج قابل ارائه (گام ۱)

گزارش کسب و کار مورد ارزیابی

- اهداف استراتژیک کسب و کار سازمان واضح و روشن شود
- اهداف پیشنهاد برنامه BI تعیین شود
- بیان نیازمندی های کسب و کار (مشکلات یا فرصت های کسب و کار)
- شرح چگونگی نرم افزار BI که توسط آن نیازها برآورده خواهد شد
- شاخه هایی که به نیازهای کسب و کار اشاره نکرده و راه حل پیشنهادی BI را به کار نمی گیرند
- نتایج تجزیه و تحلیل هزینه- سود
- ارزیابی ریسک
- توصیه هایی درمورد سیستم های عملیاتی یا فرآیندها و روندهای عملیاتی کسب و کار ، برای بهبود فرآیند کسب و کار

نکته: گزارش ارزیابی همچنین باید شامل یک یا دو صفحه بررسی اجمالی اجرایی باشد که خلاصه جزئیات این گزارش را دربرمی گیرد.

نقش های درگیر در گام ۱

- نماینده کسب و کار
- حامی کسب و کار
- تحلیلگر کیفیت داده ها
- مدیر پروژه
- کارشناس موضوع

خطرات ناشی از انجام ندادن مرحله ۱

حامی قدرتمندی برای راه اندازی کسب و کار که از اهداف استراتژیک کسب و کار پشتیبانی کند ، نمی توان یافت.

گام دوم ارزشیابی زیر ساختار سازمان (Enterprise Infrastructure Evaluation)

این گام ارزشیابی امکانات موجود در سازمان و تعیین امکانات مورد نیاز برای انجام پروژه هوش تجاری است. در این گام باید برای پروژه یک برنامه‌ریزی داشت.

زیرساختارها دو دسته‌اند: فنی و غیر فنی

زیر ساختار فنی شامل موارد زیر می‌شود:

- سخت افزار
- شبکه
- میان افزار(واسط)
- سیستم‌های مدیریت پایگاه داده‌ها و ابزارها
- زیر ساختار غیر فنی عبارتند از:
- مدل منطقی داده‌ها
- متادیتاها
- رهنمون‌ها و رویه‌هایی که باید انجام شود

زیر ساختار فنی:

سخت افزار: تجهیزات و ادوات فیزیکی قابل لمس سخت افزار نامیده می‌شود.

- در حال حاضر از چه سخت افزاری استفاده می‌شود؟ به چه سخت افزاری نیازمند و قیمت آن چقدر است؟

در این بخش ارزشیابی باید شامل موارد زیر باشد:

- برنامه هایی که برای ادغام سخت افزارهای جدید با پیکربندی سخت افزاری موجود ، وجود دارد.
- سیستم های مدیریت پایگاه های داده در برنامه ی سخت افزار انتخاب شده باید دسترسی به پایگاه داده و رشد استفاده از آن را کاملاً دربر بگیرد. (مقیاس پذیری نیز یکی از مسائل اساسی است که باید به آن اشاره نمود).
- انتخاب برنامه تحت تأثیر الزامات قابلیت همکاری بین برنامه‌های سخت‌افزاری مختلفی می‌باشد. (چنانچه نیاز باشد)
- هزینه و بازگشت سرمایه (ROI) برای سه قسمت قبلی عامل های کنترل کننده محسوب می شوند.

مقیاس پذیری (Scalability): تغییر متناسب متغیرهای تابع بر حسب تغییر متغیرهای مستقل در یک فرآیند ، مقیاس پذیری آن فرآیند نامیده می شود.

مثلاً در برنامه‌های که اعداد را مرتب می کند و خروجی این برنامه اعداد مرتب شده است ، وقتی ورودی بیشتر می شود برنامه به همان اندازه باید کار بیشتر انجام دهد. مقیاس پذیر بودن یعنی چنانچه ورودی برنامه تغییر کند ، زمان خروجی نسبت به چندین برابر شدن ورودی ، خیلی بیشتر نشود. (مثلاً اگر داده‌ها ده برابر شوند ، افزایش زمان کمتر از ده برابر شود)

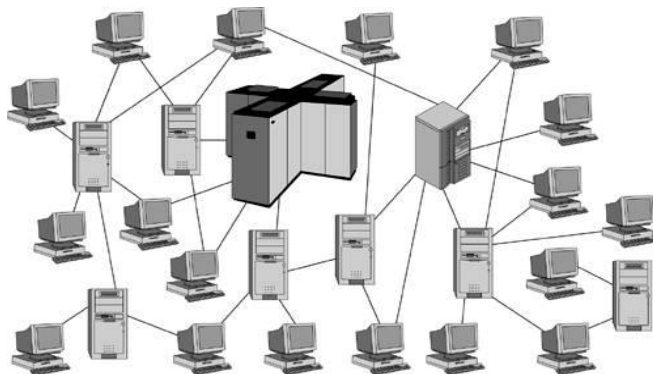
چون در پروژه‌های هوش تجاری حجم بالایی از داده‌ها وجود دارد مقیاس پذیری دارای اهمیت است در این پروژه‌ها دائماً داده‌ها آپدیت می شوند و الگوی دسترسی به داده‌ها مرتباً تغییر می کند یعنی جنس گزارشات در این پروژه‌ها فرق می کند و تعداد و تنوع گزارشات بسیار متفاوت است ، بنابراین سخت افزار موجود باید متناسب با داده‌ها مقیاس پذیر باشد.

نکته: سخت افزار ، به جهت تغییرات سریع در موارد زیر ، باید مقیاس پذیر باشد:

- حجم اطلاعات
- به روز کردن موارد روزمره
- الگوهای دسترسی به داده‌ها
- تعداد گزارشات و درخواست‌ها
- تعداد افرادی که به پایگاه‌های داده هدف هوش تجاری دسترسی دارند
- تعداد ابزارهایی که در مقابل پایگاه های داده هدف هوش تجاری به اجرا در می آیند
- سیستم های عملیاتی که اطلاعات مورد نیاز پایگاه های داده هدف هوش تجاری را مهیا می کنند

شبکه:

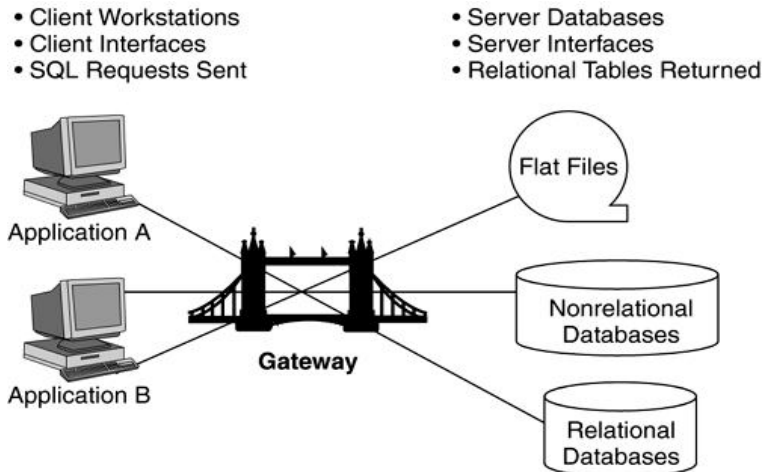
- در شبکه از چه شبکه محلی استفاده می شود؟ توپولوژی شبکه چیست؟
 - اطلاعات WAN چیست؟ آیا پهنای باند برای پروژه‌ی هوش تجاری کافی است؟
- پراکندگی کنترل شده سخت افزار



میان افزار یک لایه واسط (Gateway) بین نرم افزارهای هوش تجاری است.

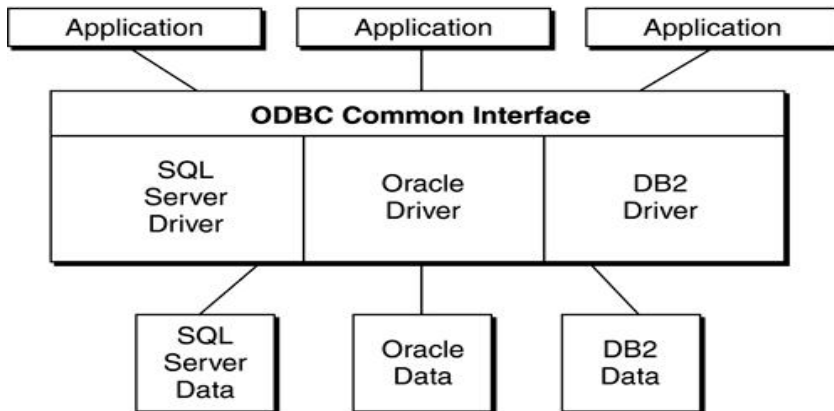
چون همه‌ی نرم افزارها زبانشان یکسان نیست و در پروژه‌ی هوش تجاری تمام نرم افزارها کنار هم قرار می گیرند نیاز مبرمی به میان افزارها است. انواع میان افزارها:

- میان افزار منطقی توزیع شده
- میان افزار مدیریت داده



درگاه های سیستم های مدیریت پایگاه های داده

- درگاه های نقطه به نقطه
- درگاه هایی که می توانند به صورت سراسری مورد استفاده قرار گیرند
- درگاه هایی که از زبان درخواست ساخت یافته (SQL) استفاده می کنند
- درگاه های مبتنی بر رابط های برنامه نویسی برنامه های کاربردی هستند



معیارهای انتخاب یک سیستم مدیریت پایگاه داده

در ارزشیابی DBMS هم همان سؤالات مطرح می‌شود ، در حال حاضر از چه DBMS استفاده می‌شود و آیا به DBMS جدیدی نیاز است؟ هزینه‌ی این DBMS چقدر است و آیا این DBMS سازگار با سیستم عامل هست یا نه؟

DBMS: در پروژه‌های هوش تجاری درمورد DBMS انتخاب شده باید ویژگی‌های زیر بررسی شود:

۱- درجه موازی سازی (درجه‌ی همسانی در بررسی درخواست‌ها و بارگذاری‌های اطلاعات):

در پروژه‌های هوش تجاری معمولاً از کامپیوترهایی استفاده می‌شود که از چندین پردازشگر بهره می‌برند. روی این کامپیوترها باید برنامه‌هایی نوشته و اجرا شوند که تکه تکه شوند و هر تکه به یک پردازنده داده شود و این پردازنده‌ها کنار هم بطور موازی برنامه را پردازش کنند.

۲- مقیاس‌پذیری پایگاه داده: چنانچه حجم داده‌ها خیلی زیاد شود عمدتاً DBMSها بسیار کند می‌شوند بنابراین باید میزان کند شدن را به نسبت تغییر حجم داده بررسی شود.

۳- قابلیت دسترسی سریع به داده‌ها: دارا بودن مکانیزمی که در حجم بالای داده‌ها بتواند به سرعت داده‌ی درخواستی را به کاربر تحویل دهد.

۴- یکپارچه سازی با اینترنت

۵- قابلیت طرح های فهرست بندی پیشرفته

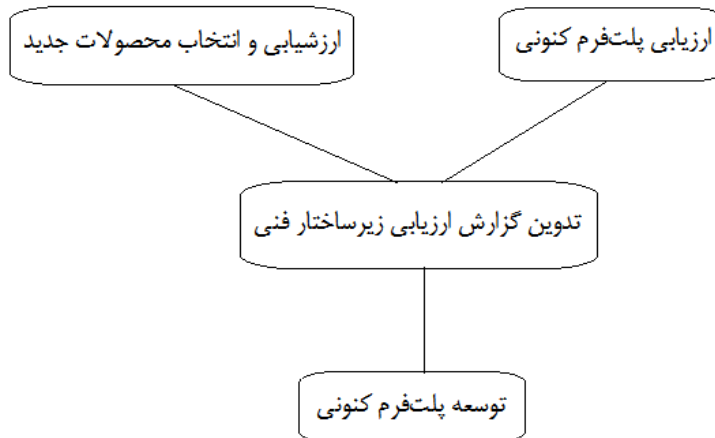
۶- تکرار برنامه های نامتقارن

۷- عملیات‌های بدون مراقبت

نتیجه گیری تفصیلی گزارش ارزیابی زیرساخت های فنی

- سرورها
- دستگاه های خدمات گیرنده
- سیستم‌های عامل
- میان افزار (به خصوص درگاه های سیستم مدیریت پایگاه داده)
- رابط های کاربری سفارشی
- اجزای شبکه و پهنای باند
- تأسیسات و قابلیت های سیستم مدیریت پایگاه داده (پشتیبان‌گیری و بازیابی ، نظارت بر عملکرد)
- ابزارهای توسعه مثل مهندسی نرم افزار کمکی رایانه و ابزارهای ETL
- ابزارهای تجزیه و تحلیل و دسترسی مثل ابزارهای OLAP و تهیه کنندگان گزارش

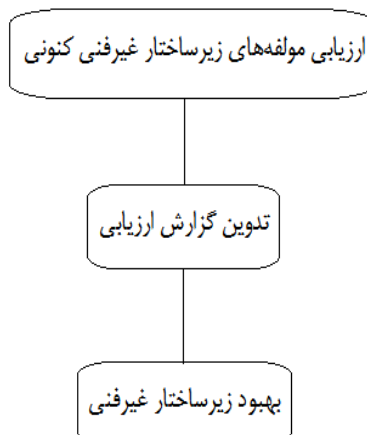
فعالیت‌هایی که باید در این گام انجام شود:



زیرساختارهای غیرفنی: عبارتند از:

- مدل منطقی داده‌ها (Logical Data Model): چارت سازمانی آن سازمان یک زیرساختار غیر فنی و مدیریتی است یا مانند نمودارهای ER در پایگاه داده‌ها (تعیین ارتباط بین عناصر).
- متادیتاها (فراداده): متادیتا به معنی اطلاعات در مورد اطلاعات است! به این صورت که هر فایلی که شامل اطلاعاتی است، توضیحی با خود به همراه دارد. متادیتا ممکن است ساختار فایل را توضیح دهد و یا ممکن است اطلاعاتی در مورد محتوای فایل داشته باشد. در واقع فراداده‌ها داده‌هایی هستند داده‌های اصلی را توصیف می‌کنند و موجب می‌شوند که داده‌های اصلی بهتر و واضح‌تر معرفی شوند. یکی از مزیت‌های متادیتا پاسخ‌گویی سریعتر به درخواست‌ها است. (صرفه‌جویی در زمان)
- رهنمون‌ها و رویه‌ها: بررسی رویه و روندهای موجود و تغییر آن‌ها در صورت لزوم.

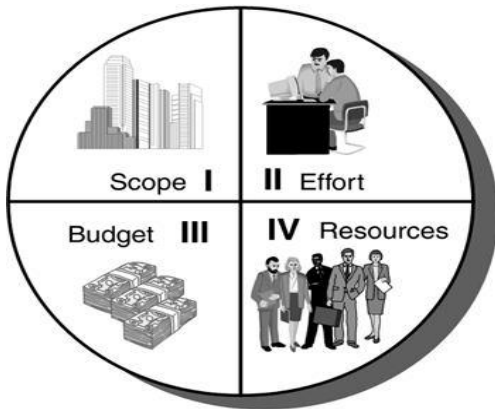
فعالیت‌هایی که باید در این گام انجام شود:



گام سوم (برنامه‌ریزی پروژه):

در این گام چهار موضوع یا چهار محدودیت پروژه باید بررسی شود:

محدودیت‌های پروژه:



Scop: تعیین محدوده‌ی حوزه‌ی فعالیت پروژه است.

یعنی اینکه در پروژه نهایتاً چه چیزی قرار است تحویل داده شود ، مثلاً برنامه‌ای تهیه شود که روی موبایل‌های اندروید قابل اجراست اما روی IOS قابل اجرا نیست.

Time: تعیین مدت زمان مورد نیاز برای انجام و زمان بندی‌های مربوط به انجام هر قسمت پروژه است.

Budget: چقدر منابع مالی موردنیاز است؟

Resource: منابع انسانی ، چه کسانی پروژه‌ی را انجام می‌دهند؟ که همان تیم پروژه هستند.

Quality: کیفیت یعنی مشخصات یک مورد یا موقعیت خاص.

مانند کیفیت یک نوت بوک: مارک ، رم ، cpu و ...

کیفیت با یک سری صفات خاصه تعریف می‌شود که برخی از این صفات خاصه کمی و برخی کیفی هستند. کمی مانند وزن نوت بوک و کیفی مانند رنگ نوت بوک.

به‌دنبال یک کیفیت خوب بودن به معنای این است یک استاندارد مشخص مورد نظر است. یعنی یک معیار یا سنجه باید برای مشخصات قرار داد. مثلاً برنامه‌ها با قابلیت مقیاس پذیری سنجش کیفیت می‌شوند. مانند ISO ها که هم سنجه‌ها هستند که باید بررسی شوند.

Preparing a project plan: خروجی گام سوم شیئی است به نام Charter ، که به معنی منشور

است. و شامل تعیین مجریان ، میزان بودجه و زمان پروژه می‌باشد. که یک قرارداد و یک آیین نامه است که تدوین شده و باید به مفاد آن احترام گذاشت.

مقدمات تدوین Charter:

ابتدا باید لیست کلیه‌ی فعالیت‌ها و وظایف و زیر وظایف تهیه شود. در این بخش باید کارها بطور ریز بینانه شکسته شوند و کلیت‌ها کنار گذاشته شوند و دلیل این کار ، ساده‌تر شدن کارهای بعدی است.

کارها چقدر زمان می‌برند ، حالا هر چقدر کارها شکسته‌تر باشند ساده‌تر می‌توان زمان بندی کرد در این بخش می‌توان تخمین زد هر وظیفه یا زیر وظیفه چقدر زمان می‌برد که می‌توان از مدیریت کنترل پروژه برای این کار استفاده کرد و روش دیگر رأی‌گیری است که می‌توان متوسط زمان را از کارشناسان این کار دریافت کرد.

فعالیت‌هایی که باید در این گام انجام شود:

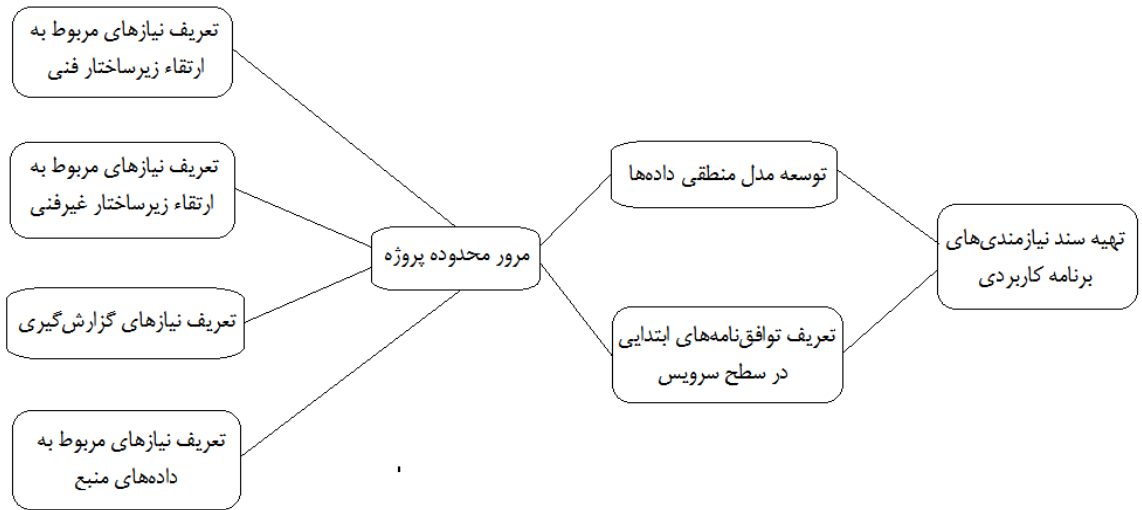


تخمین (Estimate): نزدیک بودن به واقعیت تخمین نام دارد.

این که برای انجام کار به چه منابعی نیاز است ، باید بررسی شود. این منابع و وظایف چه وابستگی به هم دارند در پروژه‌ها باید وابستگی‌ها را حتی المقدور کم کرد ولی به صفر نمی‌رسد چون پروژه به صورت تکه تکه به هم وابسته‌اند.

تشخیص مسیر بحرانی با توجه به وابستگی‌ها: مسیر بحرانی آن مسیری است که زمان پروژه را نشان می‌دهد. یعنی با توجه به مسیر بحرانی می‌توان گفت پروژه در چه مدت زمانی تمام می‌شود. متالیست‌هایی که روی مسیر بحرانی هستند اگر زودتر یا دیرتر تمام شوند روی زمان پروژه تأثیر می‌گذارند، فعالیت‌هایی که روی مسیر بحرانی هستند باید توسط افراد حرفه‌ای انجام شود زیرا اگر به تأخیر بیافتند زمان پروژه هم با تأخیر خواهد افتاد. وقتی تمام کارهای فوق انجام شود یک Plan برای پروژه وجود دارد که با جزئیات بیان شده است.

گام چهارم (تعریف ملزومات پروژه): جمع آوری فعالیت‌هایی است که در سه گام قبل انجام شد.



گام چهارم تعریف کامل‌تر و جزئی‌تر گام‌های قبل است.

در این گام باید به این سؤالات پاسخ داده شود.

سازمان به چه اطلاعاتی نیاز دارد؟ چه سؤالاتی را افراد مشغول در سازمان پرسیدند اما هنوز جواب داده نشده (نیازها)؟ داده‌هایی که برای تهیه گزارش لازم است کجاست؟ کدام گزارش‌ها مهم هستند و داده‌هایی که برای تهیه گزارش لازمند کجاست؟ گزارش‌های حیاتی و لحظه‌ای از کجا و کدام داده‌ها باید استفاده کنند؟ اگر داده‌ها آرشیو شده‌اند به اطلاعات چه مدت پیش از این نیاز است؟ داده‌ها تمیز و معتبر هستند یا نه ؟

تعریف نیازهای گزارش‌گیری: گزارش‌های حیاتی گزارش‌هایی هستند که باید لحظه‌ای آپدیت شوند. لذا

گزارش‌ها باید بسیار متنوع و شکیل برای ارائه‌ی هدف‌های مشخص به مدیر تنظیم شوند. (داشبورد ها) مثلاً در شرکت‌های زنجیره‌ای وقتی مرکز از کل شعبات یک گزارش می‌خواهد تا تمام آنها این گزارشات را تهیه و ارسال کنند حداقل یک هفته زمان لازم است و کنار هم قرار دادن این داده و Merge یا ادغام آنها با هم چند روز طول خواهد کشید. (تحلیل) یعنی حدود ۱۰ روز زودتر از تصمیم به داشتن یک گزارش مدیریتی زمان لازم است و هر یک روز تعلل در بازار جهانی موجب ضرر است.

گزارش‌هایی که معمولاً با زبان SQL نوشته نمی‌شوند گزارش‌های پیچیده هستند که توسط Data Warehousing انجام می‌شود.

گزارش گیری:

تهیه و تنظیم گزارشات خود یک هنر است معمولاً در یک برنامه گزارش‌های از پیش تعیین شده‌ای وجود دارد که برنامه‌هایی مثل حسابداری و انبار و حقوق و دستمزد یک سری گزارشات بسته‌بندی شده Canned (از پیش آماده شده) وجود دارد معمولاً این واژه به گزارش‌های از پیش تعیین شده اطلاق می‌شود اما پروژه هوش تجاری به دنبال این گزارشات نیست بلکه به دنبال گزارش‌هایی است که به مدیر تلنگری برای گرفتن گوی سبقت از رقیبان بزند.

نکته: در داده‌کاوی بخشی وجود دارد که خود سیستم به‌طور هوشمند این گزارشات را بگیرد چون وقتی یک فرد بخواهد این گزارشات را انجام دهد ذهنش یک‌جا یک وجه را بیشتر نمی‌تواند در نظر بگیرد و گزارشات معمولاً تک وجهی می‌شوند. در بخش انبارسازی ، داده‌ها به نحوی کنار هم قرار می‌گیرند که داده‌های بسیار پیچیده در کسری از زمان تحویل کاربر شوند. گزارش‌ها خروجی‌هایی هستند که به تمام افراد در سطوح کار کمک می‌کنند.

تعریف نیازهای مربوط به داده‌های منابع: تعیین سیستم‌ها و زیرسیستم‌هایی که داده‌های پروژه هوش تجاری از آن‌ها تأمین می‌شوند.

محدوده پروژه: در این گام مشخص می‌شود به چه اهدافی باید برسد و چه کمکی به سازمان می‌تواند کند.

گام پنجم (تحلیل داده‌ها) :

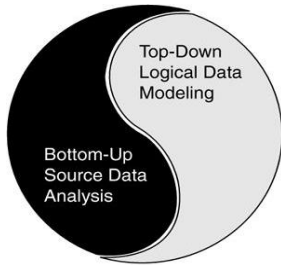
بررسی گام‌های قبلی، نیازمندی‌های جزئی و امنیتی ، پاکسازی داده و بررسی استانداردها در این گام بررسی می‌شود.

در این گام جزئیات گام چهارم بررسی می‌شود. در این گام داده‌ها باید تحلیل شوند و دقیقاً باید مشخص کرد داده‌ها از کجا و از چه محلی Source جمع‌آوری می‌شود. در این گام بیشتر در مورد کیفیت داده‌ها بررسی می‌شود و اینکه چه کسانی صاحبان داده‌ها هستند و چه روش‌هایی را برای پاکسازی داده‌ها باید استفاده کرد و اینکه آیا خطایی در داده‌ها وجود دارد یا خیر؟

این گام از دو بخش که مکمل هم هستند تشکیل شده است:

- کیفیت و استاندارد سازی (Standardization And Quality)
Bottom-up Source Data Analysis

- سازگاری و یکپارچه سازی (Integration and consistency)
Top-Down Source Data Analysis (کنار هم آوردن داده‌ها)



پاکسازی (پالایش) داده‌ها (Data Cleansing): به معنی رفع هرگونه موردی که موجب می‌شوند داده‌های موجود کثیف باشند.

مواردی که موجب کثیف بودن داده‌ها می‌شوند:

- ناموجود بودن داده‌ها: اگر مقادیر داده‌ها مشخص نباشد یا موجود نباشد ، داده‌ها کثیف هستند.

- ناسازگاری داده‌ها: اگر داده‌های یکسانی در چند جا ذخیره شده باشند و مقادیر متفاوتی داشته باشند ، ناسازگاری داده‌ها اتفاق می‌افتد و به این داده‌ها «کثیف» می‌گویند. اگر داده‌های جمع‌آوری شده از منابع مختلفی باشند ، بدون سازگار نمودن آن‌ها درد سرساز خواهند شد.

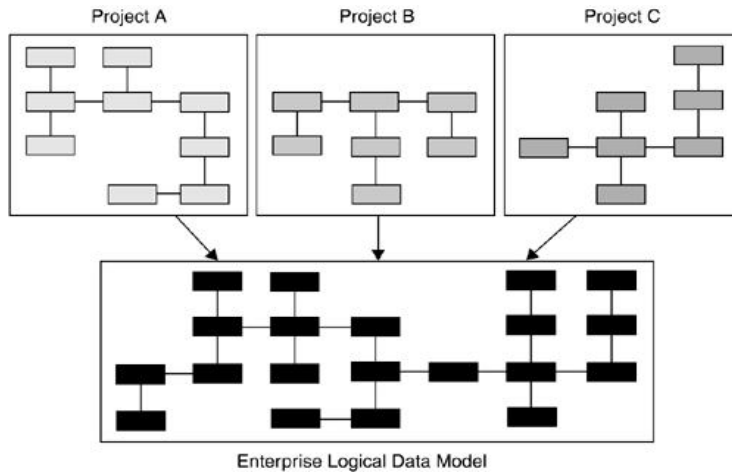
تبدیل داده (Data Transformation)(استاندارد کردن داده‌ها): برای یکسان کردن واحدهای اندازه گیری استفاده می‌شود. (مثلاً وزن شخص یکجا با kg ذخیره شده در جایی دیگر با پوند).

کاهش افزونگی داده‌ها (Data Reduction): حجم زیاد داده‌ها موجب می‌شود که پاکسازی داده‌ها با مشکل مواجه شود.

از Data Reduction وقتی داده‌ها از حجم بالایی برخوردارند استفاده می‌شود.

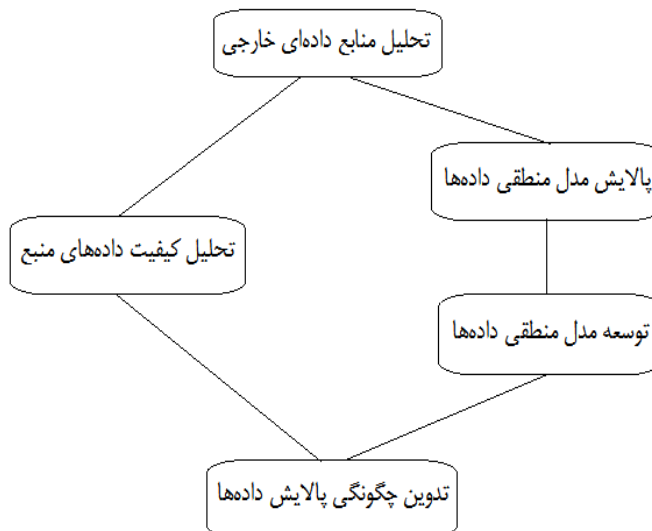
نکته: وقتی داده‌ها از جاهای مختلفی جمع‌آوری می‌شوند نباید وابسته به این باشند که از چه کامپیوتری برداشته می‌شوند یعنی نباید مهم باشد از روی چه سیستم عاملی برداشته شده و وابستگی به بانک اطلاعاتی نداشته باشد (Independence استقلال)

یک مثال تصویری برای ایجاد یک مدل داده منطقی در یک سازمان



نکته مهم: نکته مهم در گام پنجم این است که داده‌ها چگونه جمع‌آوری و تحلیل می‌شوند.

فعالیت‌هایی که باید در این گام انجام شود:



گام ششم پیشگونه سازی (Prototyping) (نمونه سازی)

Prototyping: ساختن نمونه کوچک یا ماکتی از پروژه برای ارائه‌ی بصری به سفارش دهندگان یا کاربران.

توصیه‌هایی برای ساخت Prototype:

- تعیین محدوده Prototype: باید Prototyping را کوچک ساخت تا زمان زیادی صرف آن نشود پس بعضی از موارد پروژه باید در آن قرار بگیرد.
- درک نیازهای اولیه پایگاه داده: مورد جستجو را باید شناخت و مشخص نمود که قرار است با چه دیتابیزی به انجام برسد و داده‌های درستی از این DB جمع‌آوری شود.
- انتخاب داده‌های درست: یعنی برای ساخت Prototype از بخشی از داده‌های اصلی استفاده شود و این داده‌ها باید داده‌هایی باشند که Prototype را زیر ساخت ببرد و مؤثر باشد.
- انتخاب ابزار تست: حتماً ابزارهایی باید برای آزمون و تست Prototype استفاده کرد.
- درگیر کردن مخاطبین کسب‌وکار مربوطه: افراد و کاربران واقعی پروژه باید Prototype را تست کنند تا پیشنهادات لازم را عرضه کنند.

مزیت Prototype این است که عیب سیستم مشخص می‌شود ، پس باید در اختیار کاربرانش برای تست قرار گیرد.

موارد کاربرد پیشگونه‌ها:

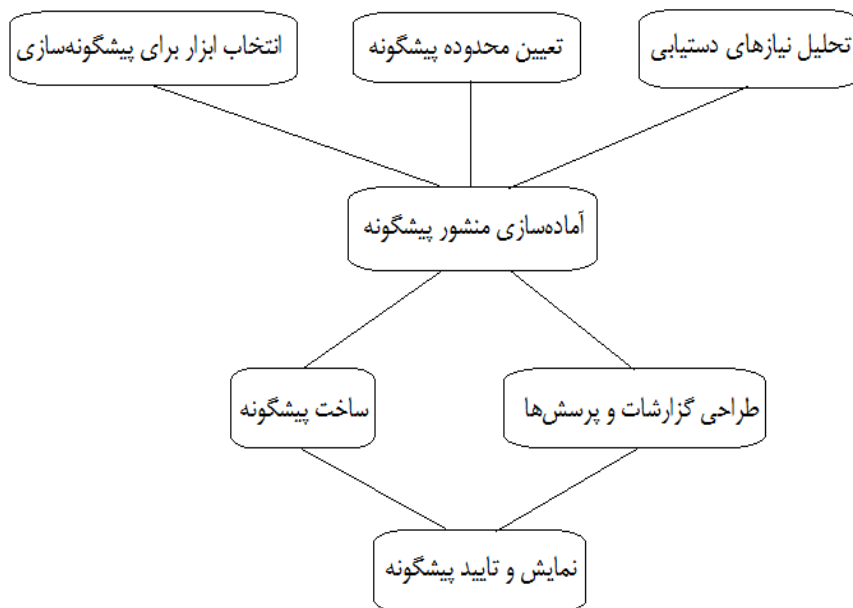
Prototype ها برای بازاریابی ، برای تهیه گزارشات و برای **User Interface** کاربرد دارند.

انواع پیشگونه سازها:

کامل‌ترین Prototype نسخه‌ی α و β برنامه است. α و β نسخه اصلی برنامه نیست بلکه نسخه آزمایشی و نوع Operational هستند یعنی برای آخرین بار تست تهیه شده‌اند و شاید دارای نقص‌هایی باشند.

فرآیند ساخت پیش‌گونه

- ۱- نمایش و معرفی نمونه اولیه: به عنوان یک نسخه‌ی نمایشی برای مدیریت و کسب‌وکار مخاطبین.
 - ۲- نمونه فانتزی اولیه: هدف درک و تجزیه و تحلیل دسترسی مورد نیاز و فعالیت‌های کسب‌وکار.
 - ۳- اثبات مفهوم نمونه اولیه: هدف ، بیان عدم قطعیت در پیاده‌سازی نمونه است.
 - ۴- طراحی بصری نمونه اولیه: درک طراحی رابط بصری و توسعه مشخصات برای رابط بصری و نمایش.
 - ۵- نسخه‌ی نمایشی نمونه اولیه:
 - انتقال چشم انداز برنامه BI به کارکنان درون سازمان یا به کاربران بیرون از سازمان.
 - تست بازار برای باقی‌ماندن یک برنامه BI در مقیاس کامل.
 - ۶- نمونه اولیه عملیاتی: ایجاد یک نمونه تقریباً کامل از برنامه‌های آلفا و بتا و در دسترس قرار دادن آن ها برای تجزیه و تحلیل بخشی از برنامه BI.
- سه چهار نوع اول را اکثر ابزارها پوشش می‌دهند و User Interface پرکاربردترین نوع Prototype است.
- فعالیت‌هایی که باید در این گام انجام شود:



گام هفتم - تحلیل مخزن متادیتا

متادیتاها (فراداده): متادیتا به معنی اطلاعات در مورد اطلاعات است. در واقع فراداده‌ها داده‌هایی هستند داده‌های اصلی را توصیف می‌کنند و موجب می‌شوند که داده‌های اصلی بهتر و واضح‌تر معرفی شوند.

این گام طراحی و توسعه مخزن متاداده است، (طراحی در گام دهم قرار دارد اما این گام‌ها به هم مربوط هستند). **تحلیل مخزن**، یعنی باید مشخص کرد که اصلاً به متادیتا نیاز است یا خیر؟ و چگونه نیاز برآورده شود. **طراحی مخزن**، که آیا متادیتا بطور متمرکز باشد یا پایگاه داده روی چند سیستم عامل قرار بگیرد، توزیع شده باشد یا نباشد. **توسعه مخزن**، نوع پیاده‌سازی که باید درباره آن هم فکر و برنامه ریزی کرد که در گام چهاردهم قرار دارد.



برای تحلیل مخزن متادیتا چالشهایی وجود دارد که مدیر پروژه باید با آن دست و پنجه نرم کند.

مثلاً برای این که این مخزن پیاده‌سازی شود نیروی انسانی کافی و مناسب در اختیار است یا خیر! این کار هزینه‌های مختص به خود را دارد زیرا این مخزن نیاز به نرم‌افزار و سخت افزار و سرورهای قوی خواهد داشت و کلاً استفاده از متادیتا برای چیست و سیاست های آن چه خواهد بود. متادیتا در واقع مانند داده‌هاست اما همان طور که داده از مدلی تبعیت می‌کند و متادیتاها هم چون داده هستند طبقاً از مدلی تبعیت خواهند کرد پس تحلیل متادیتا مانند تحلیل داده‌هاست با این تفاوت که این محتواها داده‌ها نیستند بلکه اطلاعاتی هستند که توصیفی درباره‌ی داده‌های اصلی هستند.

مزایای استفاده از متادیتا:

- سرعت پاسخ‌گویی به درخواست‌ها
- پاکسازی داده‌های کثیف (می‌توان به وسیله آن ناسازگاری و افزونگی را برطرف کرد)

گزارش زیر یک مثال کاربردی برای متادیتا است.

Monthly Sales Report				
Month	Produce	US Sales (\$)	Canada Sales (\$)	Total Sales (\$)
January	Apples	22,000	8,000	30,000
	Bananas	11,900	6,000	17,900
	Coconuts	2,000	800	2,800
February	Apples	22,500	8,500	31,000
	Bananas	10,000	5,800	15,800
	Coconuts	2,200	350	2,550
March	Apples	23,700	9,300	33,000
	Bananas	9,900	7,000	16,900
	Coconuts	2,400	750	3,150

Meta Data →

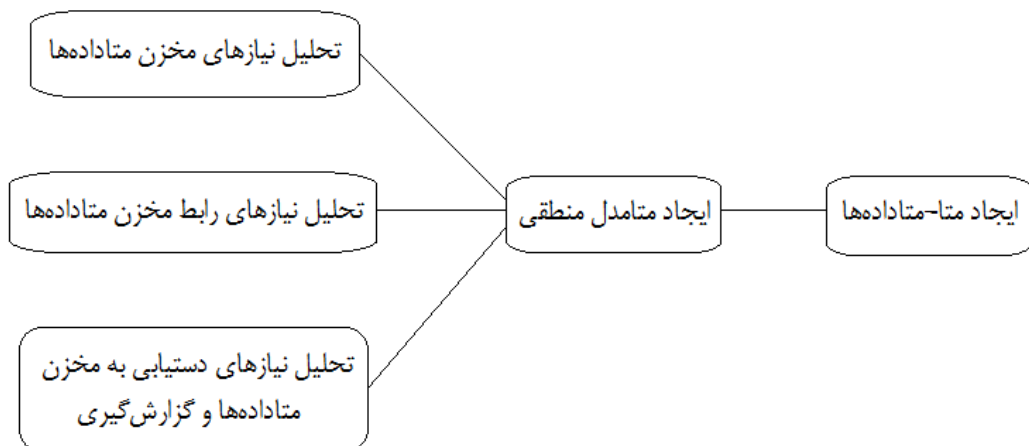
Data Quality Load Statistics:
51% of \$ values not loaded
10% of source records not loaded

فروش چند محصول در ۳ ماه از سال بررسی شده و این محصولات متفاوتند و این محصولات میوه هستند در دو جای مختلف. مثلاً در این گزارش به طور حتم رقم ۳۵۰ که کمترین مقدار جدول است به چشم مدیریت خواهد آمد و این مدیر با بردن نشانگر موس خود روی این عدد یکسری اطلاعات را در یک باکس مشاهده خواهد کرد که از مخزن متادیتا آمده است. این پیغام معناست که این عدد ۳۵۰ نهایی نیست و هنوز تمام داده‌های کامل توسط سایر شعب بدست نیامده و داده‌ها هنوز تجمیع نشده‌اند.

مثلاً هنگام تایپ متنی با نرم‌افزار Word در صورت غلط تایپ کردن واژه‌های دیکشنری Word به ما پیغام خواهد داد متن تایپ شده ی ما همان داده‌های ماست و Word با کمک متادیتا‌های خود این پیغام‌ها را برای کمک به ما خواهد داد.

در پروژه‌های هوش تجاری چون همیشه داده‌ها از منابع متفاوت و مختلفی جمع‌آوری می‌شوند وقتی کاربر سؤالی را می‌پرسد بدون وجود متادیتا آن سؤال نمی‌تواند در زمان کوتاهی پاسخ داده شود (اطلاعات تکمیلی در اسلاید).

فعالیت‌هایی که باید در این گام انجام شود:



انبارش داده‌ها (Data Warehousing)

تعریف DWhouse: فرآیند ساخت و استفاده از انبارهای داده‌ها «انبارش داده‌ها» نام دارد.

تعریف انبار داده‌ها (غیررسمی): یعنی داده‌ها از جاهای متفاوت جمع شده و در جایی نگهداری می‌شوند نحوه جمع‌آوری ، قالب، شکل نگهداری و... آنها مهم است.

تعریف انبار داده‌ها (رسمی) Data Warehouse: مجموعه‌ای از داده‌ها است که ایجاد می‌شوند تا به مدیران در فرآیند تصمیم‌سازی استراتژیک کمک کنند.

انبار داده دارای ویژگی‌های زیر است:

- **موضوع‌گرا (Subject-Oriented):** یعنی وقتی یک DW ساخته می‌شود باید دارای یک موضوع باشد و مشخص شود این DW تحت چه محوری شکل می‌گیرد. (مثل مشتری ، محصول ، تهیه‌کننده ، فروش و ...)

- **یکپارچه‌شده (Integrated):** انبار داده‌ها چون از چندین DB جمع‌آوری شده است نیاز به پالایش دارند (مثل یکسان سازی واحدها) و کنار هم یکپارچه شوند.

- **پویا پذیری (Time-Variant):** یعنی در انبار داده‌ها زمان‌های متفاوتی نگهداری می‌شود. (مثلاً داده‌های ۵ تا ۱۰ سال قبل نگهداری می‌شود). با بررسی و تحلیل داده‌های سال‌های قبل می‌توان تصمیمات استراتژیک گرفته و برنامه‌ریزی نمود.

- **پایایی (Nonvolatile):** انبار داده یک مخزن فیزیکی مجزاست. یعنی انبار داده‌ها روی حافظه‌ای ذخیره می‌شود که این حافظه دائمی است.

مثلاً در DB وقتی یک سؤال پرسیده می‌شود جوابش در یک جدول در حافظه اصلی ساخته و پس از پاسخ پاک خواهد شد.

مثلاً کارنامه دانشجویان در جدول ذخیره نخواهد شد بلکه نرم‌افزار تمام اطلاعات را در جدول جدیدی خواهد ریخت که برای پاسخ ساخته می‌شود و بعد از پرینت گرفتن کارنامه آن جدول از بین خواهد رفت.

یک پروژه‌ی هوش تجاری بدون انبار داده‌ها معنا و مفهوم ندارد و انجام شدنی نیست.

تفاوت انبار داده‌ها با DBMS ناهمگن

تفاوت دوسیستم OLAP و OLTP

برنامه‌های کاربردی پایگاه داده‌ها در دو گروه قرار می‌گیرند:

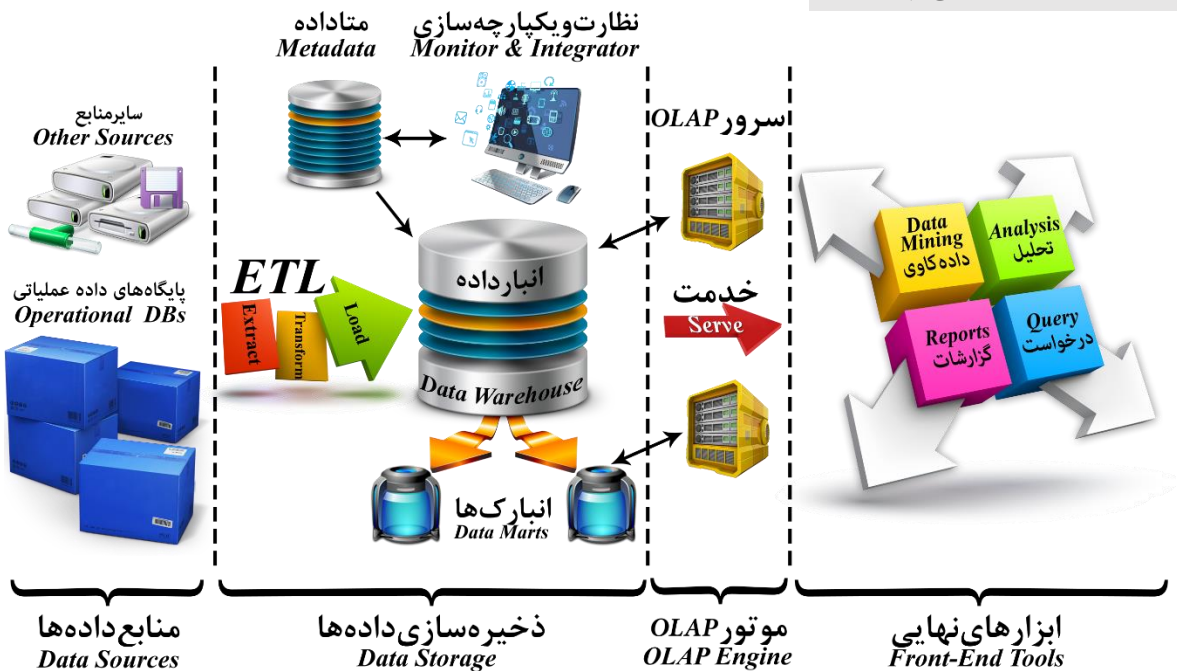
- دسته اول سیستم پردازش تراکنشی برخط **OLTP (Online Transaction Processing)** هستند یعنی کارشان پردازش تراکنش‌هاست. عملیات روزانه یک سازمان توسط این سیستم انجام می‌شود و کارت‌تحلیلی انجام نمی‌دهند و فقط گزارشات ساده و ثابت می‌دهند.

- دسته دوم سیستم پردازش تحلیلی برخط **OLAP (On-Line Analytical Processing)** یا سیستم‌های تصمیم‌یار (DSS) هستند یعنی از تصمیمات پشتیبانی می‌کنند. سیستم‌های DSS باشناسایی شاخص‌های کلیدی به مدیران کمک می‌کنند تا تصمیمات مهم و کلیدی و حیاتی بگیرند (مثلاً تصمیم‌های راهبردی مدیران فروشگاه جهت حراج‌های فصلی و ...)

تفاوت این دو نوع پردازش از زوایای مختلف در جدول زیر بررسی شده است:

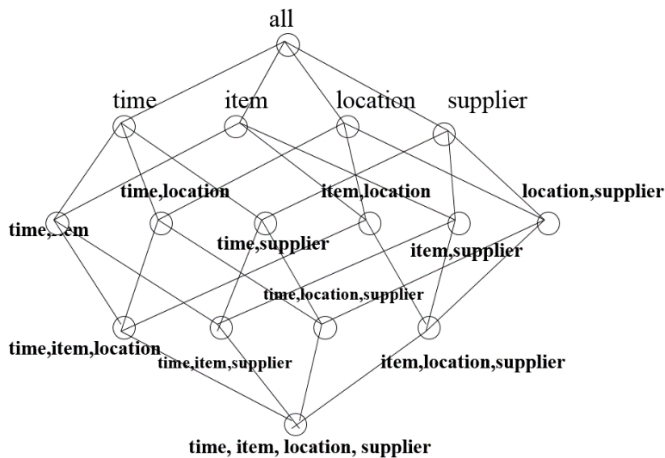
OLAP	OLTP	
(بازارگرا) مدیران و تحلیل‌گران knowledge worker	(مشتری‌گرا) کارمندان دفتری- متخصصان IT clerk, IT professional	نوع کاربران
تصمیم‌یار decision support	مبتنی بر عملیات روزانه day to day operations	عملکرد
subject-oriented	application-oriented	طراحی پایگاه‌داده
حجم بالای داده‌های وابسته به تاریخ تلخیص و یکپارچه‌سازی شده historical, summarized, multidimensional integrated, consolidated	داده‌های جاری ، جزئی current, up-to-date detailed, flat relational isolated	داده
متفاوت و موردی ad-hoc	درخواست‌ها تکراری repetitive	کاربرد
lots of scans	read/write index/hash on prim. key	دسترسی
complex query	short, simple transaction	واحد کاری
میلیون‌ها millions	ده‌ها tens	تعداد رکوردهای قابل دسترس
حداکثر صدها نفر hundreds	حداکثر هزاران نفر thousands	تعداد کاربران
100GB-TB	100MB-GB	اندازه پایگاه داده
بازده درخواست و زمان پاسخ در واحد زمان query throughput, response	بازده تراکنش در واحد زمان transaction throughput	سنجه

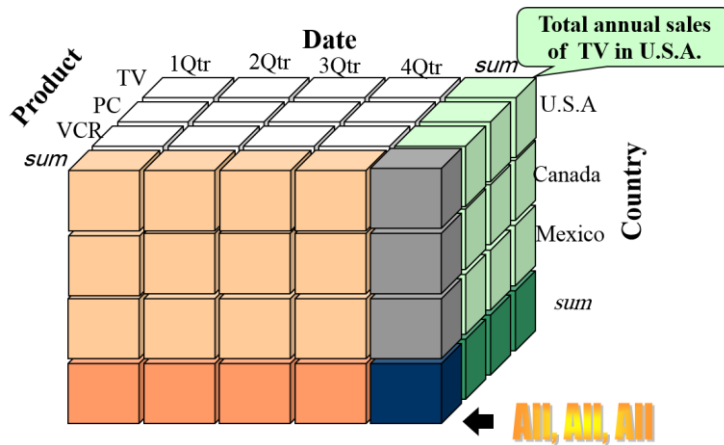
انبار داده‌ها: معماری چند لایه



مکعب: شبکه‌ی مکعب‌های کوچک (Cube: a Lattice of Cuboids)

در تحلیل‌های آماری یک سری صفات خاصه برای تحلیل استفاده می‌شود که هر کدام می‌تواند از هر جایی آمده باشد *time-item-Location-supplier* (تهیه شده- مکان- قلم- زمان) پس از چهار زاویه می‌توان داده‌ها را تحلیل نمود، این چهارصفت یک مکعب ساخته‌اند. در *OLAP*، *Cube* می‌تواند سه بعدی یا چهاربعدی یا... باشد، بُعد همان ویژگی‌هایی است که در *DWH* با آن کار می‌شود. مثلاً وقتی برای یک دانشجو ۱۰ مشخصه نگهداری می‌شود در واقع ۰ بُعد از او نگهداری می‌شود.





شکل فوق تحلیل اولیه از یک مکعب چهاربعدی است که یک بُعد زمان یک بُعد قلم جنس یک بُعد مکان و یک بُعد تهیه کننده است.

ابعاد این مکعب سه بعدی، Product-Date-Contrی، است که می‌تواند با ماتریس پیاده‌سازی شود Product یکی از ابعاد است که سه مقدار TV ، PC ، VCR دارد و Date می‌تواند چهار فصل سال باشد و Country سه کشور است این مکعب سه بعدی از مکعب‌های کوچک‌تری تشکیل شده است ۸۰ مکعب کوچک داریم و هر یک از این مکعب‌های کوچک حاوی عددی مربوط به فروش است. مثلاً در اولین مکعب بالایی سمت چپ اگر عدد ۱۰۰ باشد یعنی فروش تلویزیون در فصل بهار در کشور آمریکا و بنابراین هریک از این مکعب‌ها اطلاعات فروش را دارند. (All ، All ، All کل فروش تلویزیون در کشورهاست). این مکعب بخشی از انباره داده‌هاست.

فرض کنیم این مکعب چهاربعدی شده و بعد چهارم یعنی رنگ افزوده شود اگر سه رنگ داشته باشیم به ازای هر رنگ یک کپی گرفته می‌شود. بعد چهارم سه مقدار (سفید- مشکی و قرمز ای دارد) یک کپی برای رنگ سفید یک کپی برای رنگ مشکی و یک کپی برای رنگ قرمز. با این کار می‌توان عمل مقایسه در فروش را انجام داد و می‌توان مقدار فروش کلایی خاص با رنگی خاص در فصلی خاص و در کشوری خاص را بدست آورد.

وقتی چهار ویژگی داریم می‌توان یک مکعب چهاربعدی ساخت یا چهار مکعب سه بعدی که سطح 3-D ID است. ما به مکعب سه بعدی 3-D Cube نمی‌گویند بلکه 3-D Cube ID می‌گویند یعنی یکی از مکعب‌های سه بعدی.

Cube کل چیزی است که در صفحه می‌بینیم و Cube ID آنهایی هستند که ما درست می‌کنیم.

جمع بندی:

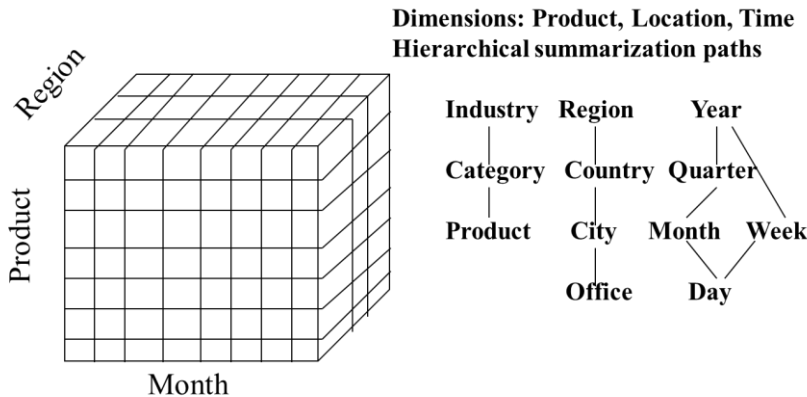
وقتی داده‌ها را از جاهای مختلف جمع‌آوری و درجایی قرار داده می‌شود برای ارتباط دادن این داده‌ها باهم از اصطلاحاتی باید استفاده کرد که یکی از آن‌ها Cube یا CubeID است.

سلسله مراتب مفهومی: اطلاعات جزئی تر یک بعد در یک مکعب ، سلسله مراتب مفهومی نامیده می شود.

مثلاً برای بیان آدرس دانشجو می توان از چندین روش استفاده کرد که تشکیل شده از شهر-استان-خیابان-کوچه و ... یعنی در سطوح مختلفی می توان آدرس را بیان کرد. مثلاً تمام دانشجویان ساکن شهر تهران-کشور-شهر-استان-خیابان-کوچه-پلاک.

مثلاً روی زمان هم می توان سلسله مراتب مفهومی تعریف کرد: سال ، ماه-هفته-روز-ساعت.

نکته: ممکن است یک بعد دارای سلسله مراتب مفهومی نباشد.



داده‌ها از سه بُعد بررسی می‌شوند اما در بُعد زمان دارای ۱۲ انتخاب هستیم و از مفهوم ماه استفاده شده است و اگر بجای ماه از فصل استفاده می‌شد دارای ۴ مقدار خواهد بود و اگر Region اگر قاره باشد حتماً ۵ مقدار خواهد داشت.

(در کتاب و اسلاید حتماً مطالعه شود چون در امتحان سؤال می‌آید)

عملیات روی مکعب‌ها

عملیات **OLAP** عملیاتی است که وقتی یک مکعب ساخته می‌شود روی مکعب می‌توان آن عملیات را انجام داد.

Roll Up: تغییر مقیاس سلسله مراتب مفهومی یک بُعد مکعب از مرتبه‌ی پائین تر (جزئی تر) به مرتبه‌ی بالاتر (کلی تر) عملیات **Roll Up** نامیده می‌شود. مثلاً تغییر از ماه به فصل یا از شهر به کشور. در یک مکعب اگر یک بعد کاهش داده شود ، این کار همان عملیات **Roll Up** است. (شکل بالا سمت راست)

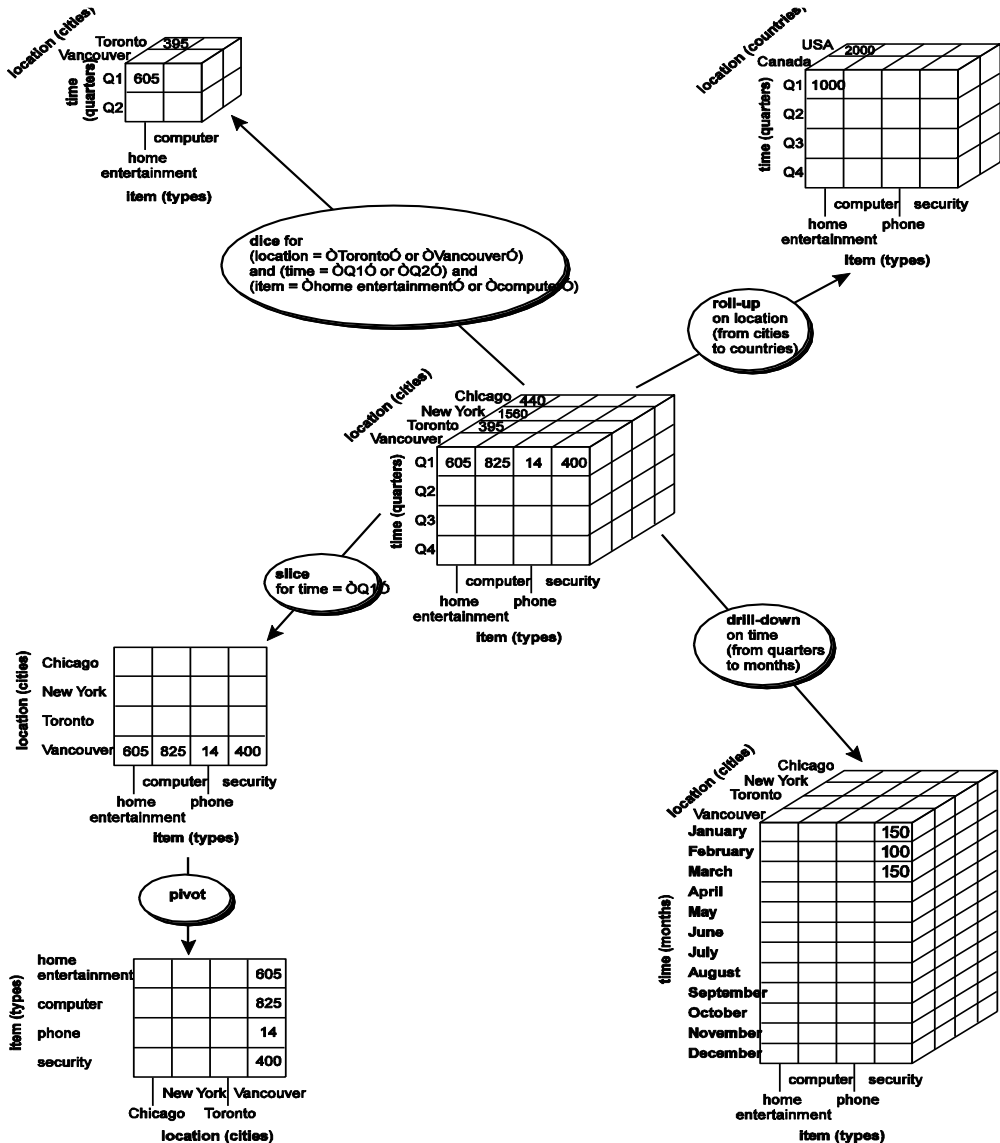
Drill Down: تغییر مقیاس سلسله مراتب مفهومی یک بُعد مکعب از مرتبه‌ی بالاتر (کلی تر) به مرتبه‌ی پائین تر (جزئی تر) عملیات **Drill down** نامیده می‌شود. مثلاً تبدیل فصل به ماه پس مکعب بزرگتر خواهد شد همچنین و چنانچه یک بُعد افزایش داده شود عمل **Drill Down** انجام شده است. (شکل پایین سمت راست)

Slice: این عمل روی یکی از بُدهای Cube یک انتخاب می‌گذارد. در این شکل روی بعد time و تایم Q1 را انتخاب کرده است یعنی لایه بالایی مکعب بالا را انتخاب کرده است. (یک بعدازسه بعدموجود) مثلاً فروش کالاها در چهار فصل در شهر Vancouver چقدر است؟

Dice: مانند Slice است اما روی بیش از یک بُعد شرط تعریف می‌کند. مثلاً در بُعد فصل ، فصل بهار و تابستان و در بُعد مکان شیکاگو و تورنتو و در بُعد کالا Entertainment و Computer.

Pivot: (چرخش روی یک بُعد) این عمل شکل نمایش داده‌ها عوض می‌کند (چیزی را جدا یا حذف نمی‌کند) مثلاً در یک نمودار دو بعدی X زمان و Y هزینه است ، اگر دوبعد با هم جابجا شوند عمل Pivot انجام شده است.

در شکل زیر ، مکعب ۳ بعدی وسط صفحه دارای سه بُعد است: Location (که شهرها مقدارش هستند) و مقدار بُعدی زمان (time) است که با توجه به چهار فصل چهار مقدار دارد و بُعد بُعدی item است.



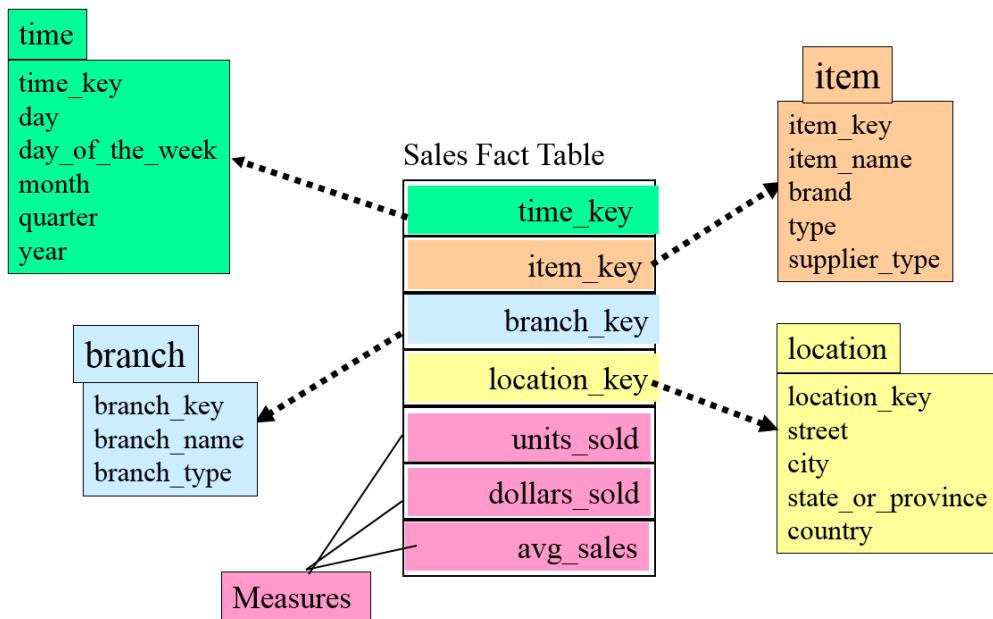
مدل سازی مفهومی از انبار داده‌ها

برای این که بتوان Cube ها را با جداول مدل سازی کرد چه باید انجام داد؟ در طراحی دیتا بیس در پایگاه داده‌های معمولی از جداول ER برای نمایش ارتباط بین موجودیت‌ها استفاده می‌شود اما در DWH برای طراحی از شماهای خاصی استفاده می‌کنند. مدل ER برای OLTP مناسب است اما برای طراحی DWH به یک شمای دیگر نیاز است. متداول ترین مدل داده‌ای برای انبار داده‌ها استفاده از جداول چند بعدی است. انواع این شماها عبارتند از:

Star -۱ ، Snow Flake -۲ ، Constellation -۳

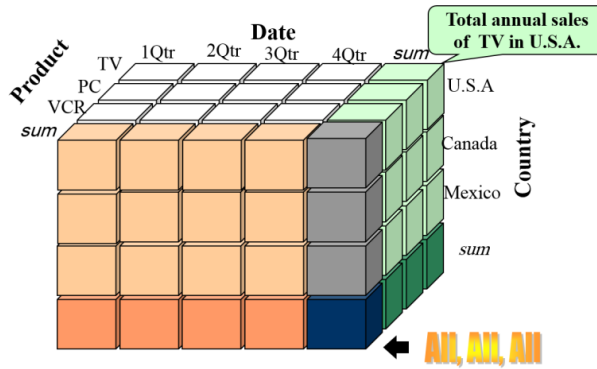
مدل ستاره (Star Schema)

در شمای Star یک جدول در وسط قرار می‌گیرد که به آن Fact Table (جدول اصلی) می‌گویند و بقیه جدول‌ها دور آن قرار می‌گیرند. در مثال زیر یک جدول در وسط و ۴ جدول در اطراف آن (جداول ابعاد) وجود دارد. به عبارتی یک Cube چهار بعدی در این مدل پیاده سازی شده است. و این چهار جدول با کمک Fact table باهم در ارتباطند.



در Fact Table هفت ویژگی وجود دارد که چهار ویژگی آن جهت ارتباط با آن چهار جدول است که همان کلیدها هستند. Branch-key ، Location-key ، item-key ، time-key و اما سه ویژگی دیگر باهم وجود دارند که کلید نیستند بلکه همان مقادیری است که در هر یک از سلول‌های مکعب نوشته شده است. (اطلاعات جمع‌آوری شده که صفات خاصه سنجه نامیده می‌شود) ، در هر سلول سه مقدار قرار دارد. به هر یک از کلیدها یک Dimension گفته می‌شود.

شکل زیر یک مکعب ۳بعدی است که ۸۰ مکعب کوچک دارد که در هر مکعب اگر از مدل Star استفاده کرد یک مقدار قرار دارد که تعداد واحد ، مبلغ واحد و میانگین فروش هستند که کلید نیستند. (کلیدها برای ارتباط با جداول است)

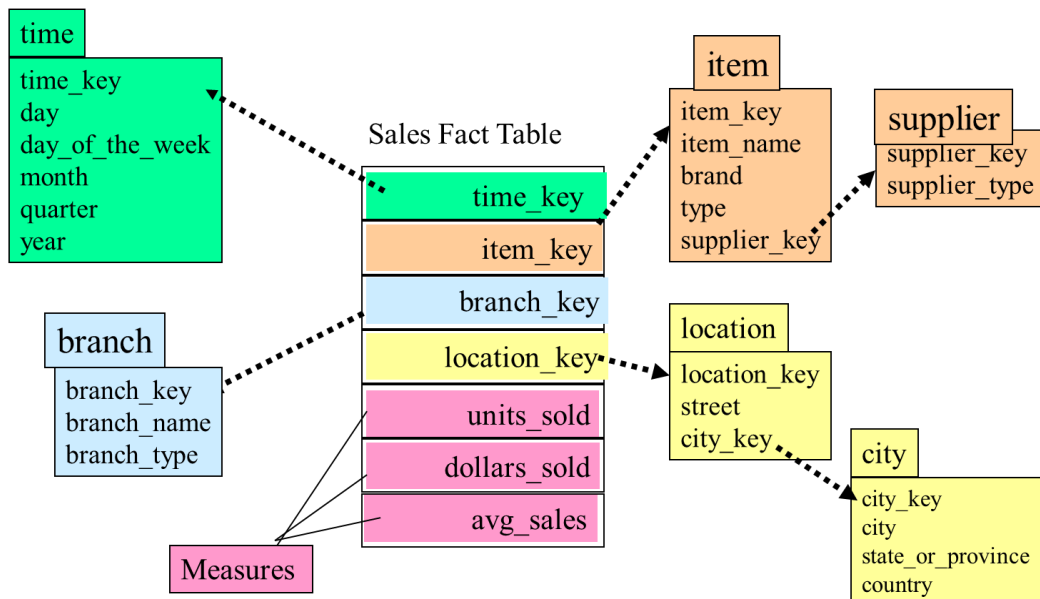


کلیدها ارتباط بین بُعدها یا جداول را مشخص می کند اما سه مقدار بعدی measure است که همان مقدار جعبه‌هاست. در مدل ستاره‌ای جداول اطراف به جایی جز Fact Table وصل نیستند.

مدل برفگونه (Snowflake Schema)

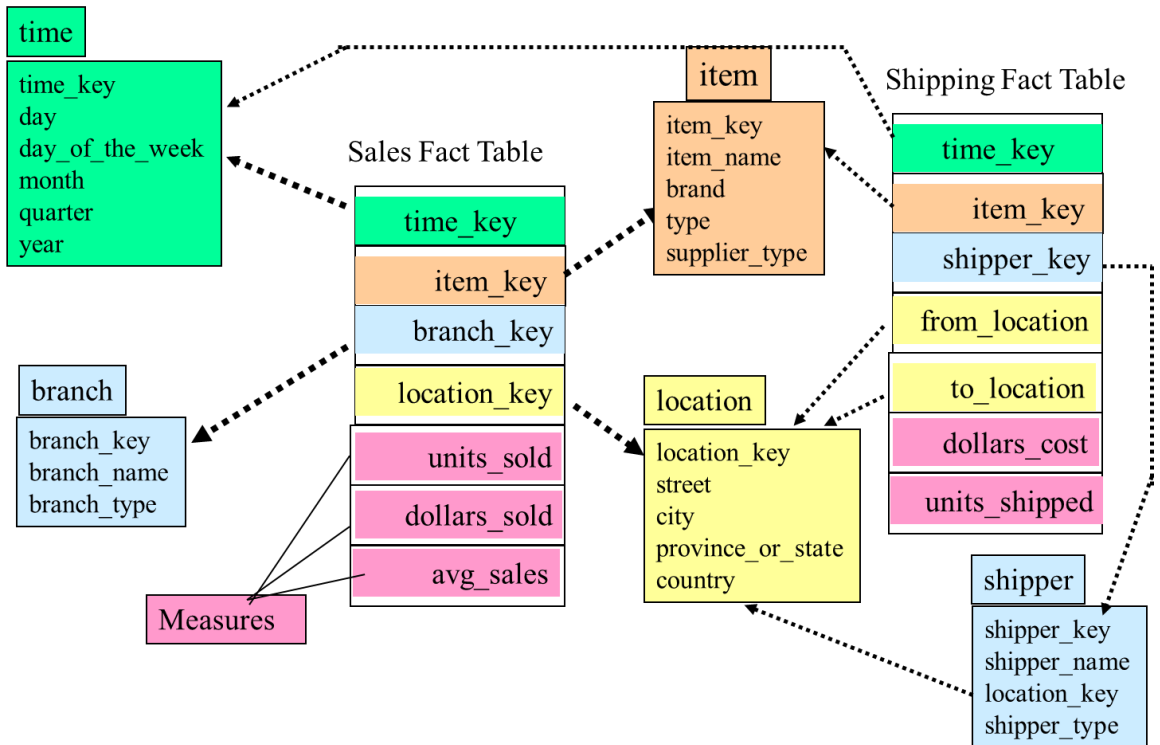
تفاوت این مدل با ستاره‌ای این است که جداول اطراف به غیر از جدول Fact به جداول دیگری متصل هستند، جدول‌های اطراف در مدل قبلی نرمال‌سازی نشده بودند اما در مدل برفگونه نرمال‌سازی شده‌اند ، در صورت وجود افزونگی در جداول اطراف جدول Fact باید جداول را نرمال‌سازی کرد در اینجا item برای نرمال شدن به دو جدول تبدیل شده است یعنی item و supplier.

این مدل از نظر تئوری بهتر است اما در مدل کاربردی از مدل ستاره‌ای استفاده می کنند چون سریع تر پیش می رود. پس افزونگی را به جان خریده و برای کمتر شدن مشکلات زمانی از ستاره‌ای استفاده می کنند.



مدل صورفلکی (Fact Constellation)

در این مدل بیش از یک جدول Fact وجود دارد اما در دو مدل قبلی تنها یک جدول Fact وجود داشت. خود این جداول Fact به جداول اطرافشان وصل هستند در اینجا برخی از جداول اطراف به هر دو جدول Fact متصل هستند. معمولاً این مدل برای سیستم‌های خیلی بزرگ استفاده می‌شود چون بسیار پیچیده اند در حالی که مدل‌های کاربردی از مدل ستاره استفاده می‌کنند. پس این مدل به ندرت استفاده خواهد شد.



معماری سرور OLAP

OLAP چندبعدی (MOLAP): سرورهای OLAP برای ذخیره‌ی مکعب‌ها از آرایه‌های چندبعدی استفاده می‌کنند ، OLAP های چندبعدی می‌گویند. مزیت این کار سرعت دسترسی بسیار بالای آن است ، چون در استفاده از آرایه‌ها سرعت دسترسی بسیار بالا است. عیب این روش این است که چون بسیاری از سلول‌های مکعب خالی هستند ، بسیاری از آرایه‌های خالی فضای خالی (Sparse) هستند. (۷۰ تا ۸۰ درصد سلول‌ها خالی هستند) لذا فضای بسیار زیادی به هدر رفته و جستجو در این ماتریس دشوار است.

به ماتریس Sparse ماتریس « خلوت یا تُتک » نیز می‌گویند.

OLAP رابطه‌ای (ROLAP): در این نوع معماری داده‌ها در یک پایگاه داده رابطه‌ای ذخیره می‌شوند. این نوع معماری در مقایسه با MOLAP از مقیاس‌پذیری بهتری برخوردار است. ولی برای پاسخ به بسیاری از Queryها کارایی کمتری دارد چون رابطه‌ای است پس از ابزار SQL استفاده می‌کند.

OLAP ترکیبی (HOLAP): این نوع معماری ترکیبی از معماری چندبعدی و رابطه‌ای است. در این نوع معماری داده‌های اصلی در پایگاه داده رابطه‌ای ذخیره شده و دیگر داده‌های تجمیعی در حافظه به صورت آرایه‌ها ذخیره می‌شوند. (مثل: SQL Server)

« این مدل از انعطاف‌پذیری بالایی برخوردار است »

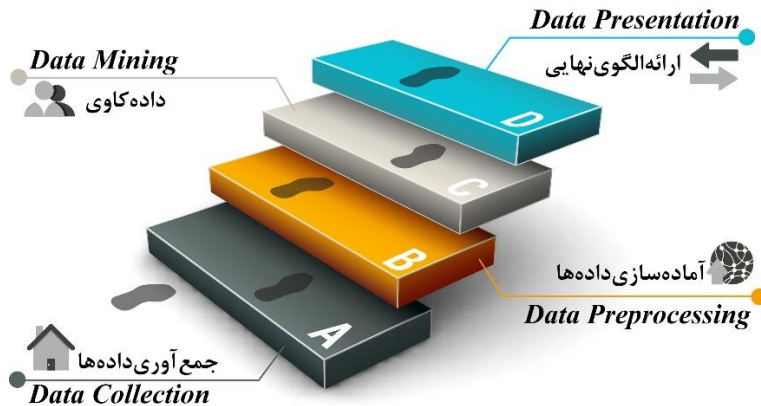
انبارک (دیتا مارت Data Mart): به زیرمجموعه‌ای از انبار داده ، انبارک یا دیتامارت می‌گویند. به عبارتی ایجاد یک انبار داده از داده‌هایی که رابطه‌ی خاصی با هم دارند و زیرمجموعه‌ای از انبار داده اصلی هستند ، یک دیتامارت می‌باشد.

مزایای ایجاد انبارک داده:

- کاهش هزینه‌های مالی و زمانی
- استفاده بهتر از داده‌ها به دلیل کوچکتر بودن حجم داده‌ها
- پاسخ‌گویی سریع‌تر به درخواست‌ها
- تغییر و به‌روز آوری سریعتر داده‌ها و متاداده‌ها
- به دلیل تمرکز نزدیک‌تر داده‌ها بر روی یک موضوع ، کاربران به زیرگروه‌های همگن تقسیم و توازن کاری ایجاد می‌شود
- حفظ امنیت فیزیکی داده‌ها (می‌توان داده‌های خاص را در دیتامارت قرار داد)
- دیتامارت‌ها برای تغذیه الگوریتم‌های داده‌کاوی مناسب هستند ، یعنی می‌توان داده‌ها را تحلیل و به وسیله آن پیش‌بینی کرد
- می‌توان دیتامارت جدید را به سادگی ایجاد کرد و با دیتامارت‌های دیگر ادغام کرد (مقیاس‌پذیر هستند)

داده کاوی

فرآیند استخراج دانش



توجه: مرحله آماده‌سازی داده‌ها حدود ۷۰ تا ۸۰ درصد کل پروژه داده کاوی را به خود اختصاص می‌دهد.

تعریف داده کاوی: استخراج اطلاعات نهان و یا الگوهای جالب و روابط مشخص در حجم زیادی از داده‌ها ، داده کاوی نام دارد.

البته می‌توان روی حجم کمی از داده‌ها نیز داده کاوی انجام داد (اطلاعات نهان چیزی است که به چشم نمی‌آید یا به ذهن مدیران خطور نمی‌کند)

معیارهای یک الگوی جالب: درست بودن - قابل فهم بودن - مفید بودن - بدیع بودن - معتبر بودن که هر یک از این خصوصیات فرمول‌های خاص خود را دارند.

مثلاً اطلاعات دانشجویان را داده کاوی می‌کنیم. یکی از الگوهای داده کاوی این است: « هر دانشجو دارای یک شماره منحصر به فرد است » اما این الگوی مفید و جدید نیست و ما خودمان از قبل این را می‌دانستیم ما باید در داده کاوی اطلاعات غیر مفید را بتوانیم حذف کنیم و نکته‌ی مهم در خروجی داده کاوی مفید بودن خروجی است.

نکته ۱: الزاماً برای داده کاوی نیاز به انبار داده‌ها نیست. اما اگر داده‌ها با انبار سازی یکپارچه شده باشند ، داده کاوی نتایج بسیار مناسب‌تری را به همراه دارد.

نکته ۲: برای داده کاوی باید داده‌های مرتبط به هم را در کنار هم داده کاوی کرد. چون از داده‌های غیر مرتبط ممکن است الگوهای استخراج شود که زیاد مطلوب نباشد.

روش‌های داده‌کاوی: روش‌های داده‌کاوی بسیار زیاد هستند و الگوریتم‌های داده‌کاوی هر کدام دسته بندی دارند. مهم‌ترین روش‌های الگوریتم‌های داده‌کاوی عبارتند از:

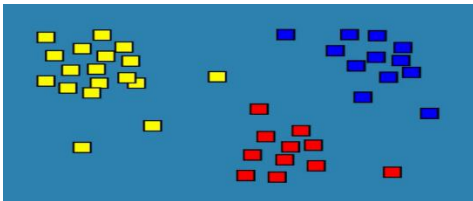


۱- **خوشه‌بندی (Clustering):** فرایند گروه‌بندی مجموعه‌ای از داده‌ها و قرار دادن آنها در طبقاتی از نمونه‌های مشابه ، خوشه‌بندی نام دارد.

یک خوشه مجموعه‌ای از داده‌هاست که نسبت به دیگر داده‌های همان خوشه شبیه بوده ولی متفاوت از نمونه‌های دیگر خوشه‌ها هستند. **خوشه‌بندی** افرادی را که به هم شباهت دارند در یک گروه قرار می‌دهد. مثلاً گروه‌بندی دانشجویان کلاس بدون در نظر گرفتن جنس یا سن خوشه‌بندی است.

« داده‌ها در خوشه‌بندی می‌توانند جنس‌های متفاوتی داشته باشند »

در بانک اسم ، فامیل ، تعداد حساب ها ، تعداد تراکنش‌ها ، موجودی فرد ، تعداد وام‌های اخذ شده برای مشتری ثبت می‌شود. حال فرض کنید برای ۱۰۰ مشتری یکی از الگوریتم‌های خوشه‌بندی اجرا شود. مثلاً الگوریتم خوشه‌بندی ۴ خوشه می‌سازد در گروه اول ۴۰ و در گروه دوم ۲۵ و در گروه سوم ۳۴ و در گروه چهارم یک مشتری قرار می‌گیرد. و تمام مشتری‌هایی که در این خوشه‌ها قرار گرفته اند به هم شبیه هستند و می‌توان از روی این گروه بندی الگویی را دریافت کرده و در اختیار رئیس قرار داد.



آن مشتری که رفتارش به هیچ مشتری‌ای شبیه نیست چون رفتارش با بقیه متفاوت است می‌تواند هکر باشد یا ممکن است آن فرد یک مشتری خیلی خوب باشد. با شناخت این فرد می‌توان سرویس‌های خاصی را به مشتریان داد.

خوشه‌بندی برای بیماران خیلی خوب است. مثلاً اطلاعات ۵۰

بیمار خوشه‌بندی می‌شوند ، و پزشک می‌تواند برای بیمارانی که در یک خوشه قرار می‌گیرند از یک روش درمانی مشابه استفاده کند.

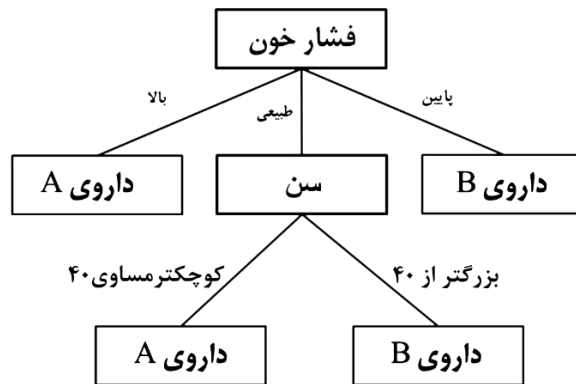
۲- طبقه‌بندی (Classification): فرآیند گروه‌بندی مجموعه‌ای از داده‌ها و قرار دادن آنها در طبقاتی مجزا براساس یک معیار یا مشخصه ویژه ، طبقه‌بندی نام دارد.

به آن مشخصه یا معیار خاص « ناظر (Supervise) » یا « برچسب (Lable) » می‌گویند.

جدول زیر مشخصات ۱۲ بیمار را نشان می‌دهد. فشار خون سه مقدار دارد: بالا ، پائین و طبیعی. اگر روی این داده‌ها داده‌کاوی انجام شود چه اطلاعاتی که به چشم نمی‌آید نتیجه می‌شود؟

ID	Sex	Age	Blood P.	Drug
1	Male	20	Normal	A
2	Female	73	Normal	B
3	Male	37	High	A
4	Male	33	Low	B
5	Female	48	High	A
6	Male	29	Normal	A
7	Female	52	Normal	B
8	Male	42	Low	B
9	Male	61	Normal	B
10	Female	30	Normal	A
11	Female	26	Low	B
12	Male	54	High	A

در این مثال داده‌ها با یکی از روش‌ها طبقه‌بندی شده‌است. یکی از این روش‌های محبوب درختان تصمیم هستند که از الگوریتم‌های طبقه‌بندی هستند و تنوع زیادی دارند.



نمونه‌ای از یک درخت تصمیم

این درخت یک مدل است که از داده‌کاوی نتیجه می‌شود.

سؤال: این مدل علاوه بر پیشگویی چه چیزی را بیان می‌کند؟

جواب: در این مدل جنسیت وجود ندارد و معنی آن این است که جنسیت برای تجویز داروی A یا B تأثیرگذار نیست.

نکته دیگری که همین مدل ساده مشخص می کند این است که مرز سنی ، ۴۰ سال است (یعنی سن بزرگتر از ۴۰ و کوچکتر یا مساوی ۴۰). مدل مشخص می کند وقتی شخص دارای فشار طبیعی است می توان با توجه به سن ، دارو را تجویز نمود.

از این مدل نتیجه می شود که رنج سنی ۴۰ سال مرز داروی A و B را مشخص می کند.

ID	Sex	Age	Blood P.	Drug
1	Male	20	Normal	A
2	Female	73	Normal	B
3	Male	37	High	A
4	Male	33	Low	B
5	Female	48	High	A
6	Male	29	Normal	A
7	Female	52	Normal	B
8	Male	42	Low	B
9	Male	61	Normal	B
10	Female	30	Normal	A
11	Female	26	Low	B
12	Male	54	High	A

برای طبقه بندی داده ها ابتدا یکی از ستون ها به عنوان ستون کلاس (یا دسته بندی) در نظر گرفته می شود.

ستون دارو در اینجا به عنوان برچسب یا Label کلاس در نظر گرفته می شود (به این دلیل که تمایل داشتیم بدانیم افرادی که دارای سن و جنسیت و فشارخون مشخص هستند ، باید چه دارویی مصرف کنند). در این مثال سن یا سایر مشخصه های جدول را هم می توان به عنوان برچسب کلاس تعیین کرد. اگر فشارخون به عنوان Label کلاس در نظر گرفته شود ، سیستم مدلی ایجاد می کند که با توجه به سن ، سن و جنسیت ، فشارخون تخمین زده می شود.

اگر ستون دارو به عنوان برچسب یا Label کلاس در نظر گرفته شود ، جنسیت ، سن و فشارخون متغیرهای مستقل و دارو متغیر وابسته می باشد.

تفاوت بین طبقه بندی و خوشه بندی

در خوشه بندی بیماران دسته بندی می شدند ، اما در طبقه بندی خودشان براساس دارو به دو دسته تقسیم شده اند. در طبقه بندی باید یک ستون از کلاس وجود داشته باشد ، مثلاً ستون دارو. (Supervise)

روش های طبقه بندی شده روش هایی با ناظر (Supervise) هستند و روش های بدون ستون و کلاس روش های خوشه بندی هستند.

سؤال: چقدر می توان به این مدل اعتماد نمود؟

پارامترها و معیارهایی برای اعتماد به مدل طراحی شده‌اند. یکی از معیارها معیار « صحت مدل » است. (Accuracy) یعنی مدل از هر صد مورد چقدر درست تشخیص می‌دهد؟ بیمارجدیدی که وارد می‌شود شاید مشخصاتش در جدول نباشد اما با کمک همین مدل‌سازی می‌توان برایش دارو تجویز نمود.

تعداد داده‌ها لزوماً نمی‌تواند در صحت مدل مؤثر باشد. (اما معمولاً داده‌های زیاد صحت بهتری دارند)

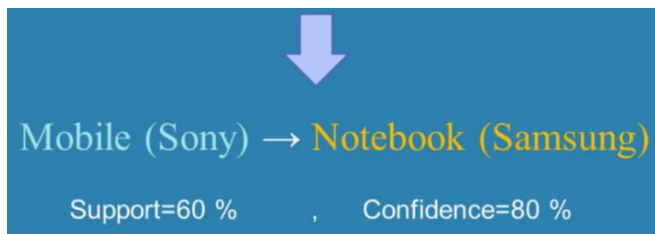
سایر معیارها مثل دقت (Precision) ، Recall ، F- Measure را می‌توان مطالعه نمود.

سؤال: تفاوت پروژه هوش تجاری با سایر پروژه‌ها مثل MIS یا EES یا IMS و ... در چیست؟

جواب: یکی از تفاوت‌ها در این است که در پروژه هوش تجاری از ابزارهای تحلیل هوشمند مثل داده‌کاوی استفاده می‌شود.

۳- قوانین (Rules): (قوانین انجمنی): در قانون ارتباط بین ستون‌ها مشخص می‌شود.

نمونه‌ای از یک قانون: ۸۰ درصد از مشتریانی که موبایل سونی می‌خرند در بازه‌ی زمانی ۶ ماه بعد کامپیوتر سامسونگ خریداری می‌کنند.



Support = 60%: یعنی ۶۰٪ افراد موبایل سونی با نوت بوک سامسونگ را با هم می‌خرند و Confidence= 80% بین کسانی که موبایل سونی خریدند (شاید بیشتر از ۶۰ درصد) ۸۰ درصدشان نوت بوک سامسونگ خریدند.

در داده‌کاوی قوانین حتماً نباید اتفاقاتی را که می‌افتد را بیان کرد بلکه می‌توان اتفاقاتی را که نمی‌افتند از دل داده‌ها بیرون کشید. به عنوان یک نتیجه احتمالی مدیر شرکت می‌تواند برای ۲۰٪ افرادی که موبایل سونی خریدند و نوت‌بوک سامسونگ نخریده‌اند ، یک کاتالوگ رایانه سامسونگ ارسال کند.

نمونه سؤال:

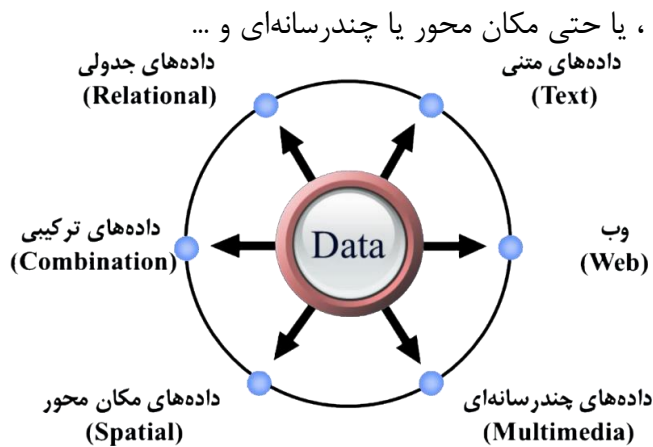
۱- تفاوت Classification با Clustering در چیست؟

۲- در Clustering چه کاری انجام می‌شود؟

۳- در Classification چه کاری انجام می‌شود؟

۴- داده‌کاوی چیست ؟ چه کاری در داده‌کاوی انجام می‌شود؟

انواع داده‌ها: می‌توان داده‌ها را روی انواع داده‌ها مشخص نمود. این داده‌ها می‌توانند مثلاً وب باشند مانند



GIS و SPS مکان محور هستند. مثلاً نقشه‌ی یک شهر را بدهیم و ببینیم در هر جا کودکستان هست در نزدیکی‌اش حتماً یک پارک هست. شرکت‌های آب و برق و گاز حتماً داده‌های مکان محور دارند.

نرم‌افزارهای داده‌کاوی:

SPSS - Celementine - WEKA - Repaid Miner - SAS Inter Prise Miner - IBM Intelligent Miner

بهترین آنها «رپید ماینر» است. چون به‌روزتر است و کاربری آن بهتر است.

چالش‌های داده‌کاوی:



- معمولاً سیستم‌های داده‌کاوی خاص وجود ندارند. مثلاً برای هر رشته‌ای نرم‌افزار خاصی وجود ندارد.

- توسعه‌ی در روش‌های مقیاس‌پذیر ، وقتی داده‌ها زیاد می‌شوند روش‌های داده‌کاوی زمان بسیار زیادی می‌خواهند.

- توسعه روش‌های تعاملی روش‌هایی هستند که کاربران هم در آن تعامل داشته باشند ، کاربران بتوانند از الگوریتم درخواست ویژه‌ای داشته باشند.

- توسعه داده‌کاوی توزیع شده ، باید روش‌های داده‌کاوی وجود داشته باشد تا از داده‌های توزیع شده پشتیبانی کنند و در حال حاضر عمومیت ندارند.

- توسعه داده‌کاوی بصری ، بسیاری از روش‌های داده‌کاوی دارای خروجی پراز فرمول‌های پیچیده است ولی انسان‌ها تمایل دارند که با خروجی همراه با تصاویر مواجه شوند.

گام هشتم - طراحی پایگاه داده هوش تجاری

در یک پروژه‌ی هوش تجاری داده‌ها سرمایه هستند زیرا یک مدیر با داده‌ها می‌تواند مدیریت کند، حال برای نگهداری این داده‌ها نیاز به جایی است که به آن پایگاه داده‌های هوش تجاری گویند. فلسفه‌ی طراحی پایگاه داده‌ی هوش تجاری با طراحی پایگاه‌های داده عملیاتی (Operational DB) متفاوت است.

Operational DB

معمولاً گزارشاتی که از این سیستم‌ها گرفته می‌شود تکراری هستند یعنی از مزیت و چارچوب خاصی برخوردارند مانند رزرو کردن بلیط هواپیما که گزارشات چارچوب داری را در پایان روزهای کاری خواهد داشت یا گزارشات بانکها یا رزرو هتل.

اما در پروژه‌های هوش تجاری گزارشات ساده نیستند یعنی با گونه‌های متفاوتی از پرسش‌ها روبرو هستیم که وقتی گونه‌های پرسش‌ها متفاوت می‌شود گزارشات آن هم متنوع می‌شود که برای تحلیل تجاری کاربرد دارند.

۱- پس اولین تفاوت این است که گزارشات پایگاه داده‌ی عملیاتی ثابت و روتین است اما در هوش تجاری روزانه نیستند و از تنوع خیلی بالایی برخوردارند.

۲- در سیستم‌هایی که از پایگاه داده عملیاتی استفاده می‌کنند زمان پاسخ به پرسش‌ها خیلی سریع است اما گزارشاتی که از هوش تجاری گرفته می‌شود به دلیل بزرگ بودن پروژه‌های هوش تجاری زمان پاسخ به این گزارشات خیلی طولانی‌تر خواهد بود و معمولاً برای جواب دادن به پرسش‌ها چون سیستم درگیر با پایگاه داده بسیار بزرگی می‌شود زمان زیادی برای محاسبات خواهد بود.

۳- معمولاً در سیستم‌های پایگاه داده عملیاتی بیشتر بر روی کاهش افزونگی تمرکز می‌شود (با نرمال سازی یا طراحی با ER). پس تفاوت دیگر این است که، پایگاه داده‌های عملیاتی تا جایی که ممکن است نرمال سازی شده‌اند و افزونگی‌هایشان کاسته شده و جامعیت را تا حد امکان دارا هستند اما در پایگاه داده های هوش تجاری برعکس است. در هوش تجاری با قرار گرفتن جداول مختلف کنار هم بدون شک نرمال سازی اتفاق نمی‌افتد چون نرمال سازی با شکستن جداول حاصل می‌شود اما در BI جداول از جاهای مختلف جمع‌آوری و در یک جدول گنجانده می‌شود.

۴- معمولاً در پایگاه داده‌های عملیاتی اطلاعات خیلی قدیمی نگهداری نمی‌شوند ، امروزه در محیط‌های کاری اطلاعات را تا جای خاصی آرشیو یا بایگانی می‌کنند و در این سیستم‌ها گزارشات خیلی قدیمی وجود ندارد و بیشتر روزانه‌اند ، اما در DB هوش تجاری از داده‌های گذشته برای تحلیل بسیار استفاده می‌شود.

۵- معمولاً داده‌ها در پایگاه داده‌های عملیاتی خلاصه سازی شده نیستند اما در پایگاه هوش تجاری اعداد و ارقام خلاصه سازی شده‌اند و به شکل تلخیص یا تجمیع هستند. مثلاً نمره دانشجویان ذخیره شده و معدل آنان ذخیره می‌شود. چون نگاه BI خیلی از بالاست داده‌ها تک به تک نیستند بلکه کلی‌نگری دارد. مدیران سطح بالا در BI نباید جزئی‌نگر باشند. سند و داده‌های یک مدیر سطح بالا باید تجمیع و تلخیص شده باشد و مدیران سطح بالا باید تفکر جمعی داشته باشند.

نکته: دیتابیس‌های هوش تجاری با یک تغییر کوچک تحت تأثیر قرار نمی‌گیرند یعنی در کسری از ثانیه آپدیت نمی‌شوند. در حالی که DB های معمولی در آن واحد آپدیت می‌شوند.

(فصل ۲ و ۳ کتاب استاد مطالعه شود ، فصل ۳ کتاب در امتحان می‌آید)

مدل‌های طراحی پایگاه داده

مدل‌های منطقی طراحی DB (Logical Data Design):

در Data WareHousing بحث شد. به سه شکل طراحی انجام می‌شود. (مثال فصل ۳ کتاب)

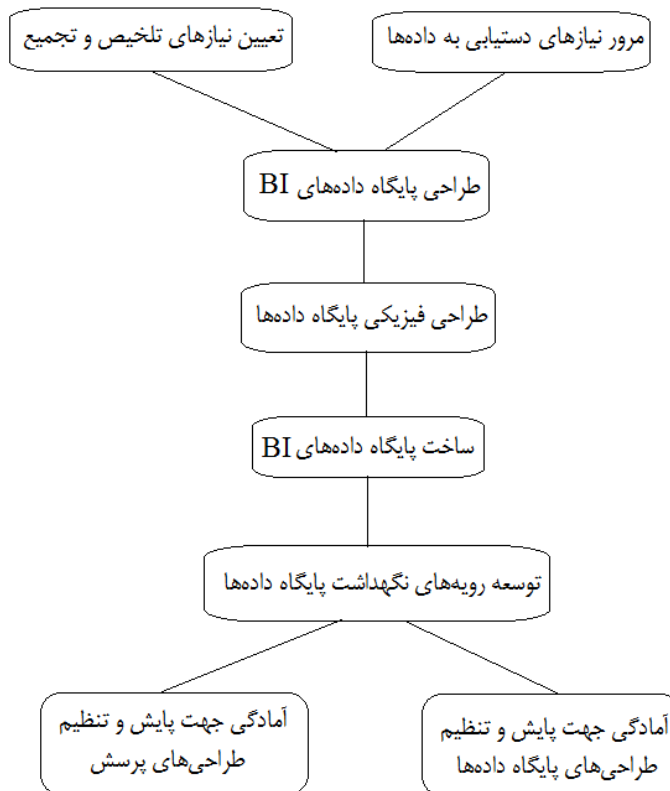
- ستاره‌ای The Star Schema
- برف‌گونه The Snowflake Schema
- صور فلکی The Constellation Schema

مدل‌های فیزیکی طراحی DB (Physical Database Design):

- Implementation Option
- Physical Dataset Placement
- بخش‌بندی (Partitioning): افراز داده‌ها روی سرورهای مختلف ، که به واسطه‌ی آن می‌توان پردازش موازی را انجام داد. لذا درجه‌ی موازی سازی بالا می‌رود.
- Clustering: یعنی کاری انجام گیرد تا داده‌ها توزیع شوند و به دلیل زیادی کاربران سرعت پاسخ دهی پایین نیاید.
- شاخص‌گذاری (Indexing): شاخص‌ها ساختمان داده‌ای هستند که با کمک آن ساختار می‌توان به سرعت به آن داده‌ها دست یافت. (برای دسترسی سریع به داده‌ها)
- Reorganisations

- Backup And Recovery: پروژه‌های هوش تجاری حجیم هستند لذا Backup گرفتن داده‌های آن طبق روال ویژه‌ای باید انجام شده و Recovery به صورت اتوماتیک انجام شود.
- اجرای موازی کوئری‌ها (Parallel Query Execution): چون در داده‌کاوی برخی از کارها مدت زمان بسیار طولانی (مثلاً سه چهار ماه یا سال) طول می‌کشد ، از سیستم‌هایی می‌توان استفاده کرد که چندین پردازنده (۲۰ ، ۳۰ ، ۱۰۰ پردازنده) داشته باشد ، پردازش را می‌توان بین آنها تقسیم کرد.

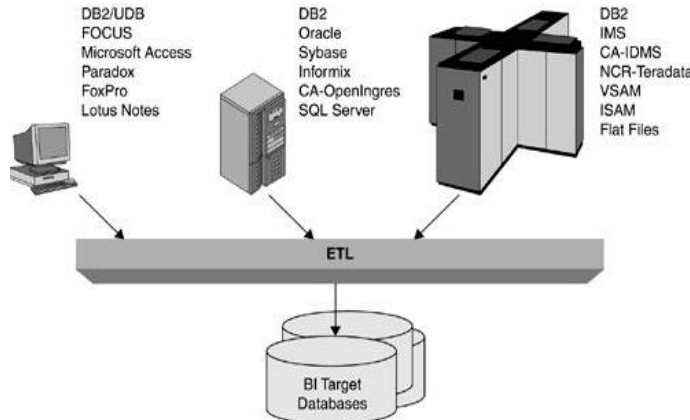
فعالیت‌هایی که باید در گام هشتم باید انجام شوند عبارتند از:



گام نهم - طراحی ETL (Extract/Transform/Load)

ETL داده‌های مختلف را از منابع مختلف جمع‌آوری نموده و با یک قالب استاندارد در پایگاه داده هوش تجاری ذخیره می‌کند.

ETL از سه گام تشکیل شده است.



- ۱- استخراج (Extract): در این گام مشخص می‌شود منابعی که داده‌ها از آن‌ها جمع‌آوری شده‌اند ، چه منابعی هستند و داده‌ها از چه منابعی به‌دست آمده‌اند.
- ۲- تبدیل (Transform): داده‌های متنوع برای کنار هم قرار گرفتن نیاز به تبدیل دارند این داده‌ها دارای قالب یکسانی نیستند و واحدهای متفاوتی دارند (مانند مقیاس قد و وزن افراد) حالا این داده‌های متنوع ممکن است از DBها و سیستم عامل‌های مختلفی جمع‌آوری شده باشند.
- ۳- بارگذاری (Load): داده‌ها را بر روی حافظه فیزیکی ذخیره می‌کند.

در این گام یک قانون به نام ۸۰/۲۰ (یعنی ۸۰٪ و ۲۰٪) وجود دارد در این گام ۸۰٪ کارها Transform

است یعنی از سه کار ETL، ۸۰٪ کار روی تبدیل داده‌هاست. (فصل ۲ کتاب داده کاوی دارای تکنیک هایی برای تبدیل داده‌هاست مطالعه شود)

برای آماده سازی داده‌ها در T چه کارهایی می‌توان انجام داد؟

نبودن مقدار برای داده: برخی سیستم‌ها بعضی از داده‌هایشان مقدار ندارد مثلاً یک بیمار در داده‌های بیمارستان وزن نداشته باشد یا مثلاً آدرس یک دانشجو وجود نداشته باشد، حال برای تحلیل این داده‌ها یا باید از آن‌ها صرف نظر کرد یا باید برای آن‌ها تعداد در نظر گرفت.

داده‌های پرت: داده‌هایی هستند که از نظر مقدار درست می‌باشند اما با محدوده‌ی اکثر داده‌های موجود تطابق ندارند. مثلاً می‌خواهیم بدانیم مردان ۳۰-۳۵ چه سائیزی برای کفششان استفاده می‌کنند (این داده ها رنج ۴۲-۴۱ دارند) اما وجود یک داده مانند ۴۹ با این که واقعی است اما پرت است و برای میانگین گرفتن کار ما را سخت خواهد کرد و معدل را بالا خواهد برد.

داده‌های نويز (اشتباه): داده‌هایی هستند که ناشی از اشتباه کاربران می‌باشند و باید تشخیص و حذف شوند. مثلاً بعد از جمع‌آوری داده‌ها یک جا شماره دانشجویی کلید است اما با ID اما جای دیگر با ST کلید شده است وقتی این اطلاعات کنار هم قرار گرفتند هوش تجاری باید متوجه باشد که این دو به یک چیز اشاره دارد.

تغییر شکل داده‌ها: مثلاً برای مقایسه دو دانشجو بر اساس قد و معدل برای پی بردن به شباهت آدم‌ها مثلاً معدل یکی ۱۷ و دیگری ۱۸ است اما قد یکی ۱۷۲ و دیگری ۱۷۵ است (معدل ۱ واحد و قد ۳ واحد تفاوت دارد) حال ۱ واحد تفاوت در معدل ارزش بیشتری دارد یا ۳ واحد تفاوت در قد، که با تحلیل مشخص می‌شود که ۱ واحد معدل تفاوت زیادی است تا ۳ واحد در قد.

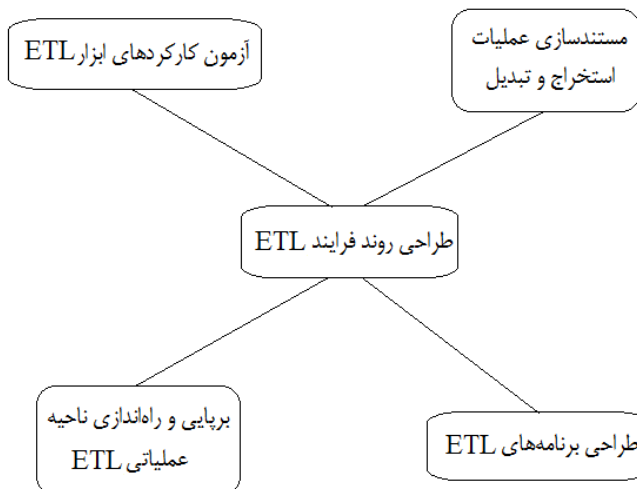
در Transform داده‌ها را نرمال‌سازی می‌کنند یعنی همه‌ی آن‌ها را در یک رنج خاص قرار می‌دهند. (معدل ۲۰-۰ را به ۱-۰ و قد ۲۵۰-۰ را به ۱-۰ تبدیل کرده و بعداً با هم مقایسه می‌کنند) (فصل ۲)

کاهش ابعاد داده‌ها: گاهی داده‌ها آنقدر حجیم هستند که نمی‌توان از همه داده‌ها استفاده نمود و برخی ستون‌ها اضافه هستند. لذا داده‌های زیادی وجود دارد که همه‌ی آن‌ها برای تحلیل مناسب نیستند (مثلاً شماره موبایل یا تلفن دانشجو) ، بنابراین باید داده‌ها را کاهش داد.

مثلاً داده‌های موجود در حد ترابایت است و زمان زیادی برای تحلیل وجود ندارد اینجا از Sampling یا نمونه‌گیری استفاده می‌شود. (تکنیک‌های آماده‌سازی داده‌ها برای تحلیل - فصل ۲ کتاب)

نکته ۱: بعد از جمع‌آوری داده‌ها اگر آن‌ها تمیز و برای تحلیل آماده شوند ، حدود ۹۰-۸۰ درصد کار پیش رفته است.

نکته ۲: Extract هم تا حدودی جزء مراحل آماده‌سازی داده‌ها می‌باشد. محصولات ETL بصورت بخش بخش هم فروش می‌رود ولی سهم عظیم کار در T انجام می‌شود.



گام دهم - طراحی مخزن متادیتا

متادیتا اطلاعاتی در مورد داده‌های اصلی بود ، دلیل استفاده از آنها وجود حجم بالایی از داده‌ها است و سیستم برای بدست آوردن جواب بدون کمک متادیتا ناتوان خواهد بود.

انگیزه‌ی اصلی متادیتا را نباید فراموش کرد هرگاه داده‌ها حجم بسیار بالایی داشته باشند معمولاً از متادیتا استفاده می‌شود. در پروژه‌های هوش تجاری برخی گام‌ها بدون وجود متادیتا امکان پذیر نمی‌باشد. مثلاً پاکسازی داده‌ها، رفع ناسازگاری داده‌ها، دیتامینینگ بدون وجود متادیتا امکان پذیر نیست ، وقتی در دیتامینینگ داده کاوی روی داده‌ها انجام می‌شود تفسیر خروجی و این که معنی این خروجی چه می‌باشد بدون متادیتا شدنی نخواهد بود. وقتی داده‌ها عوض می‌شوند متادیتا هم آپدیت می‌شود.

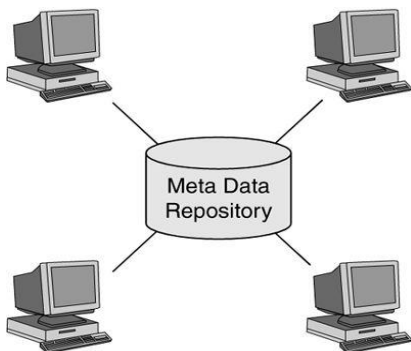
در این گام تصمیم گرفته می‌شود که متادیتا توسط خود سازمان ایجاد شود یا از منبع دیگری خریداری شود.

طراحی مخزن متادیتا: در این طراحی همان راه‌هایی را که برای طراحی مخزن پایگاه داده استفاده می‌شود ، به کار می‌گیرند یعنی دو روش: Object-Oriented-Design یا Entity-Relation-Design (ER)

چون متادیتا هم همان داده‌هایی هستند که باید در جایی ذخیره شوند پس ساختاری شبیه پایگاه داده‌ها دارند و برای طراحی مخزن آن‌ها از همان روش‌های طراحی پایگاه داده استفاده می‌شود.

طراحی و معماری مخزن متادیتا

راه حل اول استفاده از مخزن متمرکز:

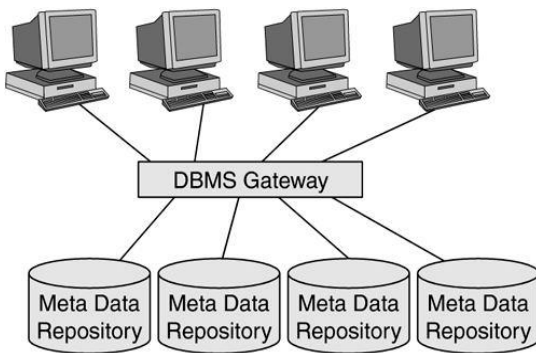


یعنی تمام اطلاعات متادیتا در یک جا به صورت متمرکز ذخیره شود. در این روش یک سرور را به متادیتا تخصیص می‌دهند و این سرویس عملاً کارش سرویس دهنده به متادیتا خواهد شد. ارتباطات و اتصالات این سرور با شبکه خیلی مهم است چون در هر لحظه که کاربر در شبکه نیاز به استفاده از مخزن متادیتا داشته باشد باید این اتصالات دارای پهنای باند مناسبی باشند تا کاربر دچار مشکل نشود. این رویکرد یک رویکرد ساده و رایج است و در شبکه‌های کوچک رویکرد مناسبی است.

اما در جاهایی که متادیتا از حجم بسیار بالایی برخوردار است خیلی مناسب نمی‌باشد چون:

- ۱- در این رویکرد ، مخزن متمرکز یک گلوگاه است ، یعنی چنانچه این مخزن متادیتا قطع شود دیگر سرویس ارائه نمی‌شود و سیستم امن نیست.
- ۲- اگر کاربران یا کامپیوترهای زیادی با این متادیتا کار داشته باشند ، ناگهان حجم ترافیک بالایی به این بخش از شبکه هجوم می‌برند و عملاً متادیتا غیرقابل استفاده خواهد شد.

راه حل دوم استفاده از مخزن غیر متمرکز: این جا مخازن متادیتا دیگر یکی نیست برای این که مشکل گلوگاه و ترافیک حل شود مخزن متادیتا به جای یکی چندتا در نظر گرفته می شود اما کامپیوترهای متصل

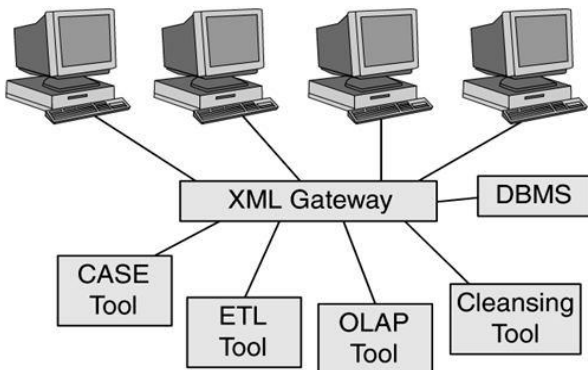


به آنها توسط یک دروازه یا Gateway به آنها وصل خواهند شد کار دروازه راهنمایی درخواست کاربران هست چون وقتی متادیتا روی چند سرور تقسیم می شود دروازه با توجه به درخواست کاربر متوجه می شود او به کدام بخش از متادیتا نیاز دارد و برای درخواست کاربر مسیریابی انجام می دهد پس Gateway سیستم مدیریت پایگاه داده ها این مدیریت استفاده از کامپیوترها را مدیریت می کند.

اما امکان دارد این سرور از نظر فیزیکی یک جا گذاشته شود چون از واژه ی توزیع شده استفاده نکرده بلکه از واژه ی غیر متمرکز استفاده شده یعنی این چهار دیسک موجود در شکل ممکن است در یک اتاق سازمان قرار بگیرد و شاید لازم نباشد در چند جای مختلف توزیع شود یا حتی ممکن است یک سرور یا دیسک در نظر گرفته شده و با استفاده از تکنولوژی خواهیم توانست آن را با روش های NM و Raid به چند بخش تقسیم کنیم.

تفاوت بین توزیع شدگی و غیر متمرکز بودن این است که حداقل نیاز برای توزیع شدگی در واقع توزیع جغرافیایی است یعنی قرار گرفتن داده ها روی سرورها و دیسک های مختلفی که در جاهای مختلفی قرار دارند.

راه حل سوم استفاده از مخزن توزیع شده: این روش در مخازن متادیتا بسیار مرسوم است. در این روش متاداده ها به شکل توزیع و مبتنی بر Xml هستند و متادیتا براساس ارتباط با ابزارهای مختلف توزیع



می شوند مثلاً متادیتای دیتاماینینگ در یک جا ذخیره می شود و متادیتای OLAP یا متادیتای Case tool و ETL هر کدام در جایی ذخیره می شوند. پس هر جا داده وجود دارد ، متادیتا نیز وجود دارد (ولی لازم نیست هر جا داده ها هستند متادیتایشان هم همان جا ذخیره شود) مثلاً در سیستم آموزش اطلاعات دانشجویان را داریم اما

چون اطلاعات ما به قدر کافی حجیم نیست پس متادیتاهای این داده ها در سیستم آموزش نگهداری نمی شوند. پس توصیف داده ها همان متادیتاها هستند اما دلیلی وجود ندارد که ما کنار این داده ها حتماً این توصیفات را نگهداری کنیم.

معیارهای خرید یک مخزن متادیتا:

در مراحل ۱۰ به بعد در پروژه هوش تجاری یا می توان هر چیزی را خرید یا می توان آن را خود نوشت. متادیتا ریپوزیتوری هم جزء مواردی است که می توان خریداری کرد. یکی از راه حل هایی که برای انتخاب یک راه حل از بین چند راه حل پیشنهاد می شود تعیین معیارهایی جهت انتخاب است. حالا از بین این معیارها باید امتیاز دهی کرد.

هفت معیار برای انتخاب یک مخزن متادیتا وجود دارد.

معیار ۱: تطبیق محصولاتی که می خواهیم خریداری کنیم با مدل منطقی داده های ماست یعنی اگر با ER یا هر چیزی جهت انتخاب محصول اقدام کردیم این کار چقدر به مدل منطقی داده های ما نزدیک است؟

معیار ۲: محصولی را که می خواهیم خریداری کنیم آیا نیازمندی های اصلی و مهم ما را پوشش می دهد؟

معیار ۳: آیا انتخاب های اختیاری در اختیار ما می گذارد؟ یعنی امکان دارد ما بخشی از آن محصول را فقط خریداری کنیم؟ و یا بعد از خرید فقط بخشی از آن را جهت استفاده فعال کنیم و بخش های دیگر را فعال نکنیم.

معیار ۴: قابلیت توسعه پذیری دارد. یعنی این محصول در آینده قابل تعمیم و گسترش خواهد بود و می توان چیزی را به آن اضافه کرد.

معیار ۵: این محصول واسط کاربری خاصی دارد و آیا نیاز به آموزش دارد و یا نه کاربران به راحتی قادرند با آن کار کنند. (خیلی مهم)

معیار ۶: آیا این محصولات قابلیت کار در شبکه های کامپیوتری را دارند.

معیار ۷: آیا این محصول API است یا نه.

در جدول صفحه بعد چند محصول معروف مخزن متادیتا در دنیا را نام برده است. با توجه به وزن هایی که ستون آخر نشان داده است به آن ها امتیاز داده ایم مثلاً شماره ۱ یک بسیار مهم است اما شماره ۳ از اهمیت کمتری برخوردار است. پس اگر معیار اول را مخزن متادیتا برآورده کند امتیاز بیشتری خواهد داشت چون وزنش بیشتر است.

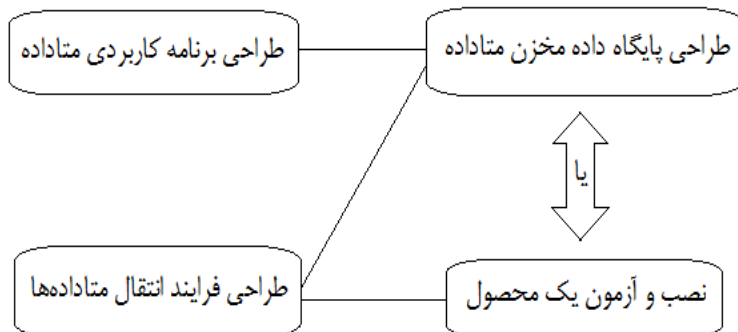
Product	Criterion #						
	1	2	3	4	5	6	7
Autumn Dictionary	3	7	6	0	9	9	10
Helixor	9	0	2	8	6	0	5
Leeches Repository	6	6	1	6	4	0	7
Springrep	8	2	0	10	10	2	10
Tamroller MDR	7	5	5	0	6	2	7

مثلاً محصول شماره‌ی ۱ Autumn Dictionary از معیار شماره‌ی ۱، ۳ امتیاز و از معیار شماره‌ی ۲ هفت امتیاز و الی آخر دریافت کرده مثلاً در معیار چهارم چون هیچ امتیازی نگرفته پس نشان از ضعف این محصول در این معیار است.

معیارهای ارزشیابی نرم‌افزاری در نهایت موضوعی می‌شوند یعنی یک جایی سلیقه و تجربه‌ی افراد هم در امتیازدهی دخیل می‌شوند.

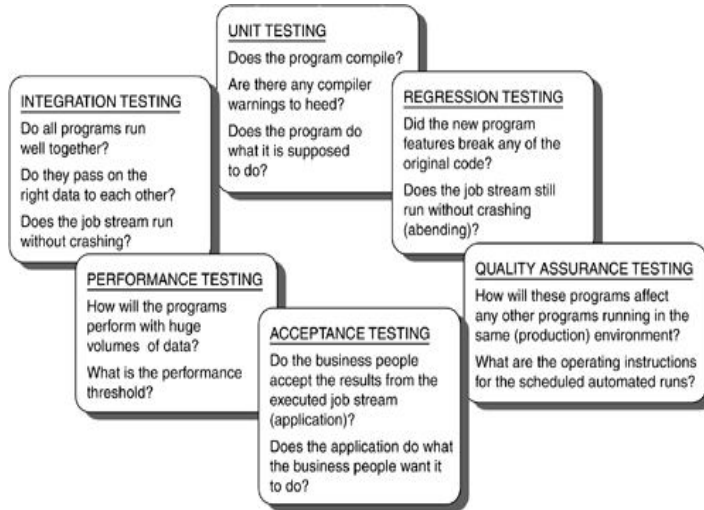
یک راه ارزشیابی مقایسه‌ی دو نرم‌افزار یا چند نرم‌افزار در یک معیار مشخص است. یعنی با توجه به توانایی‌های آن نرم‌افزارها در آن معیار و مقایسه‌ی آن‌ها با هم ما می‌توانیم در معیار واحدی آن‌ها را با هم مقایسه و به یک نتیجه‌ی مطلوب برای انتخاب آن‌ها جهت خرید برسیم.

فعالیت‌هایی که در این گام باید انجام شود:



گام یازدهم - توسعه‌ی ETL (پیاده‌سازی)

ETL بخش مهمی از یک پروژه‌ی هوش تجاری است و در ۱۶ گام پروژه حدوداً ۳ یا ۴ گام آن درگیر بحث ETL هستند در این گام تمام تقسیماتی که برای ETL گرفته شد حالا پیاده سازی می‌شود. یکی از فازهای هر پروژه ای تست است، در این صفحه از اسلاید ۶ نوع تست بیان شده است ولی شاید محصول آنقدر بزرگ نباشد که هر شش تست روی آن انجام شود اما به هر حال در پروژه‌ی هوش تجاری این ۶ تست انجام می‌شود:



Unit Testing: آزمون تک تک اجزای برنامه‌هاست. مثلاً یک برنامه‌ی ۲۰۰۰ خطی بخش‌های کوچکی از آن در این بخش تست می‌شود یعنی در واقع برنامه کامپایل می‌شود.

Integration Testing: بخش‌ها وقتی به صورت تک تک اجرا می‌شوند اما در مسیر اجرایی برنامه به صورت یکی در می‌آیند در این جا هم باید به صورت پیوسته به هم تست شوند.

بعضی از این تست‌ها روی بعضی از گام‌ها اصلاً لازم نیست مثلاً در گام‌های آخر در ۱۶ گام پروژه‌ی هوش تجاری بعضی از این تست‌ها باید روی توسعه‌ی نرم‌افزارها و سخت افزارها وجود داشته باشد. در آزمون و تست سخت افزارها و نرم‌افزارها روش‌های زیادی وجود دارند. این آزمون‌ها توسط تسترها انجام می‌شود.

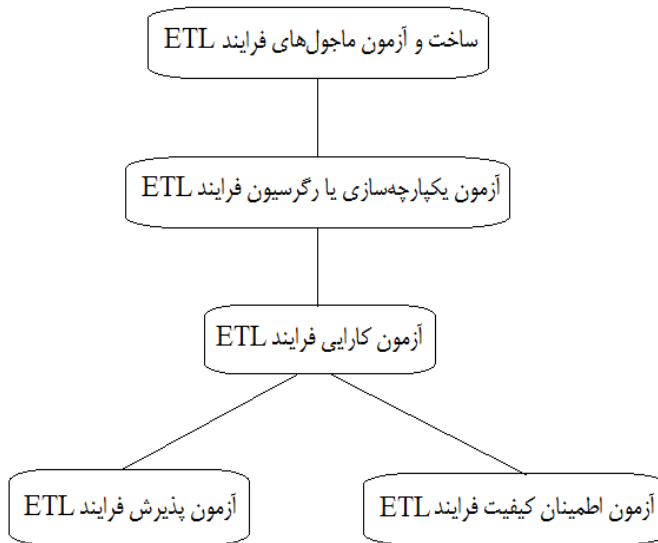
Performance Testing: در این نوع آزمایش مشخص می‌شود که در هنگام افزایش حجم داده‌ها عملکرد اجزای ETL به صورت مجزا و باهم ، چگونه خواهد بود.

Regression Testing: برای اطمینان از این که تغییرات در برنامه‌های جاری باعث ایجاد خطاهایی می‌شود یا خیر ، استفاده می‌شود.

Quality Assurance Testing: (آزمون اطمینان کیفیت) به بررسی کیفیت محصول می‌پردازد.

Acceptance Testing: (پذیرش توسط کاربران) آزمون نهایی است. در پایان کار محصول یا خدمات نهایی در اختیار کاربر نهایی قرار می‌گیرد تا مشخص شود تا چه اندازه مورد پذیرش آنها واقع می‌شود.

اقداماتی که باید در گام یازدهم انجام شود:



گام دوازدهم - توسعه‌ی برنامه‌ی کاربردی

یکی از اهداف اصلی یک سیستم تصمیم‌یار (پروژه‌ی هوش تجاری) این است که دسترسی سریع و ساده و به‌موقعی به داده‌ها وجود داشته باشد.

یعنی به درخواست‌ها سریع و به شکل‌های زیبایی (مانند داشبوردها) پاسخ داده می‌شود، معمولاً هم درصد بسیار بالایی از نیازهای کاربران از قبل مشخص است در قالب گزارشات معرفی شده‌اند. اگر نتوان این هدف را برآورده کرد ، پروژه‌ی BI موفقیت آمیز نخواهد بود و با شکست مواجه می‌شود.

به‌طور غیررسمی: گام دوازدهم مرز بین بحث‌های تئوری در یازدهم گام قبل و اجرایی شدن گام‌های بعد از آن است.

در گام دوازدهم:

- نیازهای نهایی پروژه تعیین و طراحی برنامه‌های کاربردی انجام می‌شود.
- چارچوب و قالب گزارشات نهایی و پرسش‌ها بررسی می‌شود.
- رابط گرافیکی کاربر طراحی می‌شود.
- طراحی یک راهنمای برخط بررسی می‌شود.
- موارد گام‌های قبل مثل آزمون‌ها بررسی می‌شود.
- پس از این موارد باید ماژول‌های برنامه‌های کاربردی اسکریپت نویسی شود.

یک برنامه‌ی معمولی وقتی نوشته می‌شود ، یا وارد کد نویسی می‌شود یا در حال پیاده سازی است ، یعنی برای این برنامه‌ها دو محیط وجود دارد که می‌توان آن برنامه را در آن محیط‌ها توسعه داد.

محیط‌های توسعه برنامه کاربردی

سازمان‌های کوچک دارای فقط دو محیط برای توسعه‌ی برنامه‌های کاربردی هستند (توسعه و تولید) ، و سازمان‌های بزرگ معمولاً دارای پنج محیط برای توسعه برنامه کاربردی هستند ، (یعنی در این پروژه‌های BI تنها کد نویسی و اجرای برنامه نیست):

- ۱- محیط پیش‌گونه سازی: آزمایش فناوری و تثبیت الزاماتی که در پروژه رخ می‌دهد.
- ۲- محیط توسعه: برنامه‌ها و اسکریپت‌ها توسط توسعه دهندگان نوشته شده و تست می‌شوند.
- ۳- محیط اطمینان کیفیت (QA): کارکنان عملیات تست برنامه‌های نهایی و اسکریپت‌ها را انجام می‌دهند قبل از این‌که به آنها اجازه داده شود که به محیط تولید منتقل شوند.
- ۴- محیط تولید: برنامه‌ها و اسکریپت‌ها پس از آغاز مرحله‌ی کدنویسی اجرا می‌شوند.
- ۵- محیط وب: تحلیل فرآیند کلیک کردن.

اقداماتی که باید در گام دوازدهم انجام شود:



گام سیزدهم - داده کاوی

امروزه روی کامپیوترها داده‌های زیادی وجود دارد در این کامپیوترها نرم‌افزارهایی وجود دارد که گزارشات ثابتی برای کاربر ارائه می‌دهند اما برخی نرم‌افزارها پا را فراتر گذاشته و با آن‌ها کاربر می‌تواند گزارشات را آن‌گونه که می‌خواهد از کامپیوتر استخراج کند.

ممکن است در داده‌ها چیزی باشد که به چشم نیاید و ممکن است الگویی باشد که مشخص نباشد چه سؤالی می‌توان پرسید تا این گونه الگوها را بتوان استخراج نمود. کار داده کاوی استخراج این الگوهاست. شاید در یک تعریف غیررسمی کار داده کاوی استخراج Patternهاست که از درون داده‌ها بیرون می‌آید که کاربران به سمت آنها فکرشان سوق پیدا نمی‌کند.

تفاوت داده کاوی و تحلیل آماری:

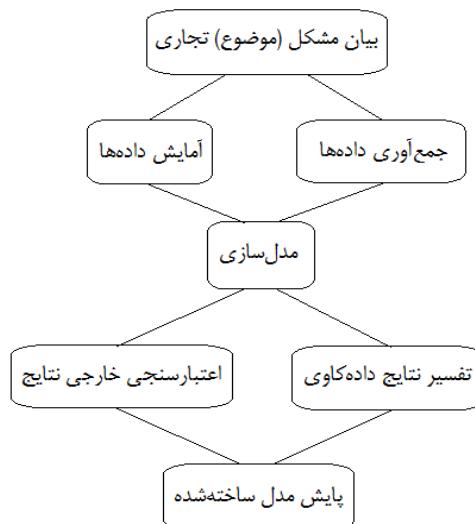
آمار چیست؟ این اولین سؤال است که تفاوتش با آمار چیست؟ یعنی آمار چه نمی‌کند که داده کاوی انجام می‌دهد؟

۱- معمولاً آمار با یک فرضیه شروع می‌شود در حالی که داده کاوی با فرضیه شروع نمی‌شود. یعنی در کار آماری باید ابتدا انتخاب کرد چه چیزی قرار است تحلیل شود اما در گام داده کاوی به فرضیه نیازی نیست و داده‌ها به کامپیوتر داده می‌شود و از درون آن‌ها اطلاعاتی استخراج می‌شوند.

۲- معمولاً فرمول‌های آماری مناسب داده‌های عددی هستند و روی داده‌های صدا و تصویر و متن خیلی قدرتمند نیستند اما داده کاوی روی داده‌های غیر ساخت یافته مانند متن بسیار قوی عمل می‌کند (به خصوص متون فارسی)

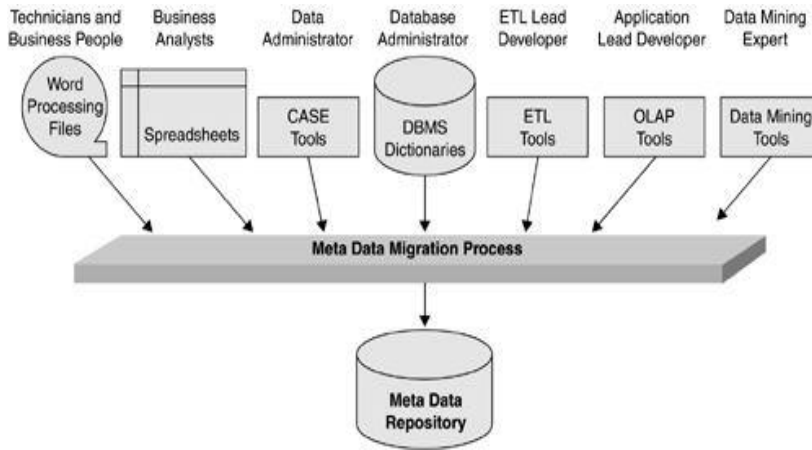
۳- خروجی‌های داده کاوی باید تحلیل شوند و نیاز به یک فرد متخصص داده کاو دارند برخی فرمول‌های آمار پیچیده است و برای جواب‌دهی زمان زیادی لازم دارند اما الگوریتم‌های داده‌ای سریع جواب‌دهی می‌کنند.

اقداماتی که باید در این گام انجام شود در نمودار زیر بیان شده است:

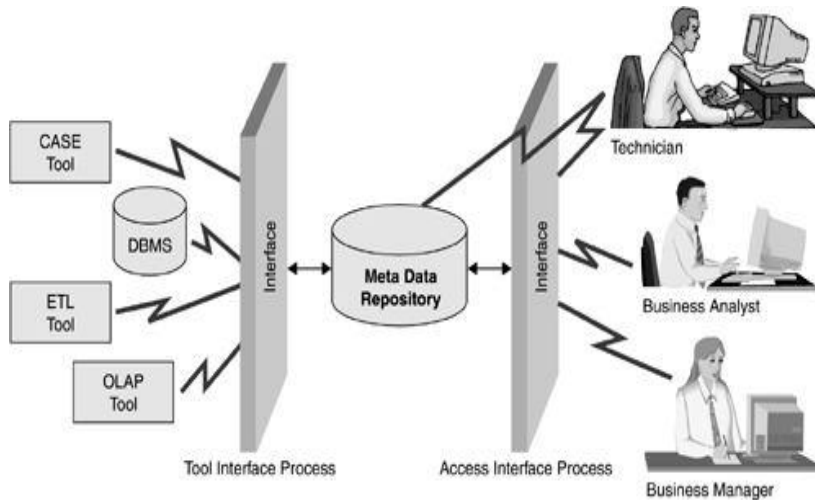


گام چهاردهم - توسعه‌ی مخزن متادیتا

داده‌های مخزن متادیتا از منابع مختلفی جمع‌آوری می‌شوند که در شکل زیر آورده شده است:



در هر مخزن متادیتا دو واسط (Interface) مجزا وجود دارد. ۱- رابط ابزار ، ۲- رابط دست‌یابی



۱- رابط ابزار: رابطی است که چگونگی دست‌یابی ابزارهای مختلف را به مخزن متادیتا کنترل می‌کند. داده‌هایی که از ابزارهای مختلف مانند: ETL Tool و OLAP Tool و DBMS و Case Tool ، در مخزن متادیتا ریخته می‌شوند ، در اینجا قرار می‌گیرد. تغییر در این ابزارها باید در مخزن متادیتا منعکس شود و این انعکاس توسط یک Interface که یک برنامه است انجام می‌شود. چون این نرم‌افزارها و ابزارها یکنواخت و یک شکل نیستند با هم سازگاری ندارد و این میان‌افزار این اطلاعات را به صورتی با هم یک جور کرده و ذخیره می‌نماید.

۲- رابط دست‌یابی: تعامل بین کاربران و مخازن متادیتا توسط این نوع Interface انجام می‌شود. کاربرانی که می‌خواهند با این مخازن کار کنند توسط این رابط با آن رابطه برقرار می‌کنند.

یکی از معیارهایی که ارزش زیادی برای خرید یک محصول این فاز دارد همین Interface است در واقع این Interface های قوی و انعطاف پذیر به ما اجازه می دهند که بخشی از آن را بتوانیم با استفاده از برنامه نویسی تغییر دهیم.

نکته: عموماً در پروژه BI از ابزارهایی استفاده می شود که برخی از آن ها در پروژه BI معروفند ، مانند: OLAP اما Case Tool ابزارهای کمکی برای کار در پروژه هوش هستند. معمولاً Case Tool یک سری ابزارهای تولید نرم افزاری به کمک کامپیوتر هستند و انواع زیادی هم دارند نرم افزارها چون تنوع زیادی دارند پس نیاز به ابزارهایی است که بتوان با آن تولید نرم افزار کرد.

تست مخزن متادیتا:

تجربه ثابت کرده است که در ازای یک روز کد نویسی ، باید سه روز تست انجام شود در اینجا هم تست وجود دارد اما فقط چهار نوع است.

Acceptance Testing ، Regression Testing ، Integration Testing ، Unit Testing همانطور که برای ابزار ETL ، شش مرحله تست گرفته شد اینجا هم چهار مرحله تست انجام خواهد شد. برنامه باید به مقدار کافی مورد اعتماد نوشته شود تا نیاز به تست های سخت نباشد ، اما کلمه « باید » معنی اش این است که اگر از روش های استاندارد مهندسی نرم افزار جهت کدنویسی استفاده نشود ، به اجبار به جای یک روز ، سه روز باید وقت صرف تست شود.

تست Regression : این Regression

ریاضی نیست در این تست بررسی می شود برنامه چقدر پوست کلفت است یعنی اگر خراب شد چقدر باعث خرابی برنامه های دیگر می شود ، به عبارت دیگر این برنامه در برابر خطاها و تغییرات چقدر به برنامه های دیگر آسیب می زند دلیل این تست در مخزن متادیتا این است که این بخش از پروژه با برنامه ها و نرم افزارهای مختلفی در ارتباط است (و با ابزارهای مختلفی کار می کند) دو مرحله آزمون Regression تست Robast بودن یا همان پوست کلفت بودن برنامه آزمایش می شود.



اقداماتی که باید در این گام انجام شود

گام پانزدهم - پیاده‌سازی

بعد از ایجاد و تست‌های متعدد برنامه‌کاربردی BI در محیط‌های مختلف ، پروژه آماده پیاده‌سازی است برای اینکه در محیط تولید قرار بگیرد تا دیگران از آن استفاده کنند.

دو روش برای پیاده‌سازی وجود دارد:

۱- پیاده‌سازی دفعی (به یکباره). معمولاً پروژه‌های هوش تجاری بسیار بزرگند بنابراین گاهی بصورت افزایشی پیاده‌سازی می‌شوند اما این کار سختی است چون نرم‌افزارها بصورت تکه تکه به سختی انجام می‌شوند.

حسن این روش این است مثلاً بعد از پیاده‌سازی پروژه شاید تعدادی داده جدید وارد پروژه شده باشد ، در روش افزایشی فقط این داده‌های جدید به محل خاص خودشان اضافه می‌شوند و نیاز نیست پروژه از ابتدا پیاده‌سازی شود در این جا نیاز نیست دیگر روی داده‌های قبلی دستکاری شود.

۲- پیاده‌سازی تدریجی. کار با یک گروه کوچکی از افراد شروع شده و به تدریج رویکرد پیاده‌سازی اجرا می‌گردد تا تغییرات لازم در سیستم ایجاد شود.

نکته: توصیه می‌شود که نرم افزارهای هوش تجاری به شکل تدریجی به سیستم وارد شود و به صورت دفعی این کار اتفاق نیفتد.

پشتیبانی و ترمیم داده‌ها:

در این گام پشتیبانی و ترمیم داده‌ها بسیار مهم است. یکی از اهداف پشتیبانی ترمیم (بازیابی) است. همیشه وقتی حجم زیادی داده وجود دارد ممکن است اتفاقات ناگواری بیفتد مثلاً تجهیزات سخت افزاری فرسوده می‌شوند یا نرم‌افزارها دچار خرابی یا ارور می‌شوند. این مشکلات داده‌ها را در معرض خطر از دست دادن قرار می‌دهند. سیستم باید بصورت اتوماتیک به وضعیت سازگار و درست قبلی برگردد یعنی بدون دخالت انسان تا حد امکان قابلیت اصلاح داشته باشد. به این مساله Recovery یا ترمیم می‌گویند. اما یکی از راه‌هایی که ترمیم سیستم را تضمین می‌کند این است که از داده‌ها نسخه پشتیبان وجود داشته باشد. در پروژه‌های هوش تجاری مهم ، حجم داده‌ها خیلی بالاست پس باید به همان اندازه هم پشتیبان داشت. کپی گرفتن این حجم بالای اطلاعات زمان بر است. بنابراین نسخه‌های پشتیبان به سه روش تهیه می‌شوند:

۱- پشتیبان‌گیری افزایشی (Incremental Backup): در این روش به تدریج و مرحله به مرحله از سیستم پشتیبان گرفته می‌شود. داده‌ها در روش افزایشی به صورت کلی بک آپ گرفته می‌شوند و در مراحل بعدی پشتیبان‌گیری فقط تغییرات را لحاظ می‌کنند. همان لحظه اول شروع سیستم و شروع رشد ، داده‌ها کپی می‌شوند و بعداً فقط تغییرات را اعمال می‌کنند.

۲- پشتیبان گیری با ساختارهای خیلی قوی و پرسرعت (High-speed mainframe backup): یا از رسانه‌های با پهنای باند بالا استفاده می‌کنند.

۳- پشتیبان گیری تکه تکه (Partial Backup): به صورت تدریجی و تکه تکه یا جزء به جزء تهیه پشتیبان از داده‌ها انجام می‌شود.

در پشتیبان گیری مشکلات زیادی وجود دارد مثلاً اگر بخواهند از ۱۰۰۰ رکورد بک آپ بگیرند وقتی از رکورد اول تا ۲۰م بک آپ گرفته شد باید از رکورد ۲۱م به بعد ادامه کار انجام شود ، اما در همین لحظه ممکن است داده‌ها توسط افراد تغییر کنند و اینگونه بک آپ سازگار نخواهد بود یعنی یکی برای ۵ دقیقه قبل و دیگری برای ۵ دقیقه بعد که داده‌هایشان با هم یکی نیست.

نظارت:

در این گام نظارت بر عملکرد منابع سیستم هم وجود دارد که از یک سری تجهیزات کامپیوتری اعم از سخت افزار و نرم افزار و کاربران یا افراد و همچنین شبکه تشکیل شده است. یعنی چگونگی کار آنها و نحوه عملکردشان کاملاً زیر نظر گرفته می‌شود. نظارت بر راندمان (نسبت مدت زمان مشغول بودن دستگاه نسبت به کل بازه‌ی زمانی مورد نظر) در سه سطح انجام می‌شود:

- نظارت بر راندمان استفاده از کامپیوتر
- نظارت بر راندمان بهره‌برداری شبکه
- نظارت بر راندمان پرسنل

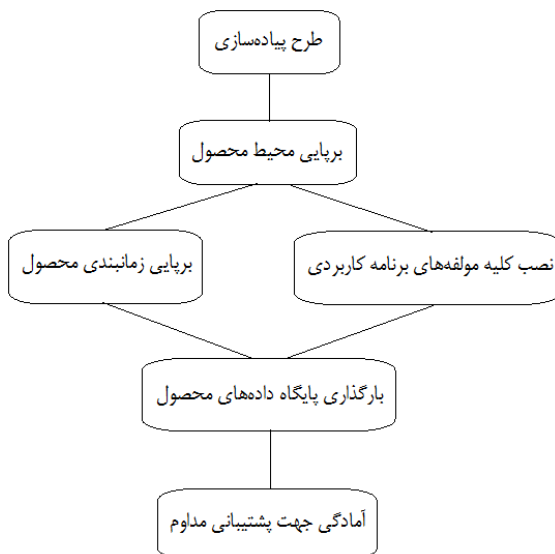
مدیریت رشد:

رشد در داده‌ها: بعد از پیاده‌سازی سیستم هوش تجاری داده‌ها رشد می‌کنند. وقتی روی جداول بسیار

بزرگ و مرکزی در دیتابیس‌هاوسینگ (Fact Table) چیزی نوشته یا حذف می‌شود Fact Table باید حتماً دوباره ریفرش شود و متادیتاها باید عوض شوند.

رشد در به کارگیری: استفاده کنندگان هم می‌توانند رشد کنند مثلاً ممکن است هنگام شروع پروژه ۱۰ نفر با آن کار می‌کنند اما با گذشت زمان افراد استفاده کننده از سیستم و داده‌ها زیادتر شوند.

رشد در سخت افزار: وقتی افراد زیاد می‌شوند سخت افزارها هم افزایش می‌یابند و این هم ما را با مشکل روبرو خواهد کرد. (مثل افزایش کامپیوترها و روترها و ...)



اقداماتی که باید در این گام انجام شود

گام شانزدهم - ارزشیابی نسخه

وقتی پروژه ساخته و تست و برنامه پیاده‌سازی شد باید مورد بازبینی قرار گیرد و نقاط قوت و ضعف آن مشخص گردد. در این گام ارزشیابی نسخه پروژه هوش تجاری صورت می‌گیرد و سؤالاتی مطرح می‌شود که مرتبط با ارزشیابی است.

سؤالات پیشنهادی پس از اجرای پروژه هوش تجاری

۱- **زمان‌بندی (Schedule):** سؤال اول این است که آیا پروژه ما به موقع تمام شده است و اگر تمام نشده دلیلش چیست؟ و چقدر دیرتر یا زودتر انجام شده است. آیا زمان‌بندی پروژه واقع بینانه بوده است؟

۲- **بودجه (Budget):** چقدر بودجه برای انجام پروژه کم داشتیم و چرا کم داشتیم در کدام گام بیشتر با کمبود بودجه مواجه بودیم؟ چه مشکلی وجود داشت و در کدام گام هزینه بیشتری شده و در کدام گام کمتر؟ و چرا؟ چطور باید برای پروژه‌های بعدی برنامه‌ریزی کرد که در بخش بودجه دچار مشکل نشویم؟

۳- رضایت‌مندی (Satisfaction):

آیا به اهداف مورد انتظار در مورد بازگشت سرمایه دست یافته‌ایم یا به آن خواهیم رسید؟ (ROI) آیا ابزارهای OLAP و سایر ابزارهای تحلیلی موجود رضایت‌مندی نیازمندی‌های تحلیلی سازمان را برآورده کرده‌اند؟ آیا این ابزارها راضی کننده هستند؟

۴- **محدوده پروژه (Scop):** آیا در طول پروژه آن چیزی که برای پروژه تعریف کردیم در همان حدود جلو رفت یا مجبور به تغییر این محدوده‌ها در گام‌های مختلف شدیم؟ آیا محدوده پروژه دچار تغییر شد؟ برای بحث و بررسی این سؤالات باید حتماً جلسه‌ای تشکیل شود و با افراد حاضر در جلسه مانند اسپانسرها، مشاورین مدیر پروژه و ... این پرسش‌ها پاسخ داده شود و بلافاصله مستند سازی شوند.

۵- **مهارت‌های مذاکره (Negotiation Skills):** آیا محدودیت‌هایی (مثل زمان ، کیفیت ، منابع و بودجه) در پروژه هست که باید در مورد آن تصمیم گیری شود؟ برای این که فرآیند مذاکره مجدد بهبود پیدا کند به چه چیزهایی نیاز است؟

۶- **رضایت کارکنان پروژه (Staffing):** چه کسانی را در این پروژه از دست داده‌اید؟ چرا؟ در کدام گام؟ آیا آنها متعلق به تیم اصلی ما بودند؟ از دست دادن آنها باعث هزینه کمتر بوده یا بیشتر؟ جایگزینی این افراد چقدر زمان برده است؟ دلایل و اثرات ترک این عضو چیست؟ آیا نقش‌ها و مسئولیت‌ها به درستی تعیین شده است؟

۷- آموزش و مهارت‌ها (Skills and Training): آیا مهارت‌های اعضای تیم پروژه برای انجام آن کافی بوده است؟ اگر خیر به چه دوره‌های آموزشی نیاز است؟ آیا آموزش‌های داده شده در قبل و حین و بعد پروژه کافی بوده است؟

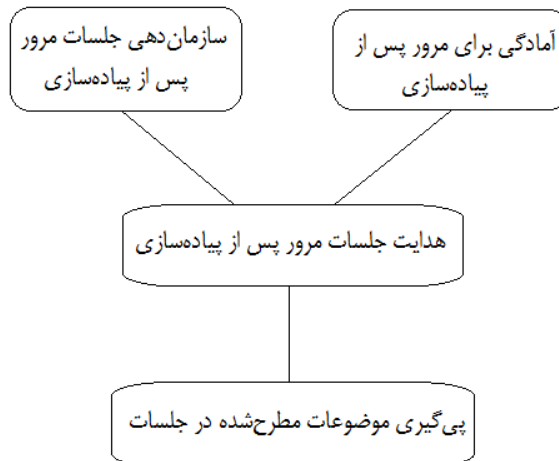
۸- گزارش‌دهی و برنامه‌ریزی پروژه (Project Planning and Reporting): آیا برنامه‌ریزی پروژه به درستی انجام گرفته است؟ آیا شکست کار درست انجام شده است؟ آیا گزارشات و قالبهای گزارش که مورد نیاز مخاطبان پروژه بوده ، تهیه و ارائه شده است؟

۹- رویکرد توسعه (Development Approach): آیا گام‌ها و فعالیت‌ها و وظایف مناسبی در مسیر ۱۵ گام پروژه انتخاب شده است؟ آیا نوع فعالیت‌هایی که برای انجام پروژه انتخاب شده‌اند انتخاب درستی بوده است؟ اگر نبودند چرا؟ آیا در بین فعالیت‌ها مورد مهمی بوده که به خوبی انجام شده یا نشده باشد؟

۱۰- مشاوران ، طرف‌های قرارداد و پشتیبان مالی (Contractors, Consultants, and Vendors): آیا ابزارهایی که برای انجام پروژه انتخاب شده‌اند مناسب بودند؟ آیا پشتیبانی ابزارها توسط شرکت‌ها و سازمان‌هایی که از طریق آنها ابزارها تهیه شده‌اند ، کافی بوده است؟ آیا سازمان طرف قرارداد ، همه مفاد قرارداد را به خوبی پشتیبانی کرده است؟

۱۱- سایر موارد (General): در دسته بندی خاصی گنجانده نمی‌شود و بسته به پروژه هوش تجاری برای یک سیستم خاص ممکن است سؤالات خاصی هم پرسیده شود که در این بخش آورده می‌شود. مثلاً: افرادی که درگیر انجام این پروژه بودند (مثل تیم اصلی و فرعی) ، آیا همیشه در دسترس بودند؟ آیا ارتباط با آنها به صورت مؤثر انجام گرفته است؟

اقداماتی که باید در این گام انجام شود در نمودار زیر بیان شده است:



شاخص‌های کلیدی عملکرد (KPI)

کارت‌های امتیاز متوازن (Balanced Scorecard)

کارت امتیازی متوازن (BSC) یک ابزار مدیریتی برای اجرای استراتژی است؛ گزارش ساختارپندی شده و نیمه استاندارد که به مدیران اجازه می‌دهد بتوانند به راحتی بر روند اجرای فعالیت‌ها توسط کارمندان نظارت داشته باشند و نتایج این فعالیت‌ها را بررسی و کنترل کنند. ویژگی اصلی کارت امتیازی متوازن در اختیار گذاردن بستری مناسب برای شناخت قوانین و روابط علت و معلولی حاکم بر دنیای کسب و کار و همچنین استخراج برنامه‌های عملیاتی برای اجرایی کردن استراتژی سازمان است.

ایده اولیه آن سال ۱۹۹۲، در خلال تحقیقات رابرت کاپلان و دیوید نورتون، در زمینه روش‌های نوین سنجش عملکرد سازمان‌ها شکل گرفت. این ایده در طول زمان توسعه و تکامل فراوان یافت تا جایی که از یک ابزار سنجش و اندازه‌گیری عملکرد، اکنون به یک نظام مدیریت استراتژیک تبدیل شده است. کارت امتیازی متوازن استراتژی سازمان را از چهار جنبه کلیدی «مالی»، «مشتریان»، «فرایندهای داخلی» و «رشد و یادگیری» بررسی و صورت‌بندی می‌کند.

در کتاب آقای پارمنتر دو زاویه دید دیگر هم به این موارد اضافه شد که از ۶ زاویه یا

پرسپکتیو یا منظر به این سنجش‌ها نگاه می‌کند.

دو منظری که بعداً اضافه شده و در اکثر کتابها نیست:

Environment/Community (محیط جامعه) و Employee Satisfaction (رضایت کارکنان)

روش کار به این صورت است که ابتدا در وجه مسایل استراتژیک، موضوع‌های استراتژیک تعیین می‌شود. سپس برای شناخت دقیق‌تر نحوه امکان دستیابی به اهداف مورد نظر در موضوع‌های استراتژیک، به سراغ وجه فرایند و عملیات رفته، مشخص می‌کنیم که برای برآورده ساختن انتظارهای ذی‌نفعان و تحقق موضوع‌های استراتژیک، فرایندهای داخلی سازمان باید چه شرایط و ویژگی‌هایی داشته باشند و در این حوزه باید به چه اهدافی دست یابیم. در نهایت در وجه رشد و یادگیری به شناسایی سرمایه‌گذاری‌های لازم بر روی منابع انسانی، سیستم‌های اطلاعاتی و فرهنگ سازمانی می‌پردازیم تا امکان و بستر دستیابی به اهداف تعیین‌شده در وجه فرایندهای داخلی فراهم شود.

خلق کنندگان کارت امتیازی متوازن چهار منظر (به اضافه دو منظر جدید) یا وجه اساسی را برای آن مطرح نمودند. این منظرها، سازمان را از کسب موفقیت یا شکست، آگاه می‌نماید. وظایف زیر با نگاه به این منظرهای ذکر شده، ارائه می‌گردد:

۱- **منظر مالی (Financial):** یعنی از منظر مالی به شرکت نگاه کنیم استفاده از دارایی‌ها و به کارگیری بهینه از سرمایه‌ها و هرچه بحث مالی در آن وجود دارد به عبارتی سیستم از منظر مالی بررسی می‌شود.

۲- **منظر مشتری (Customer Focus):** وقتی یک سازمان مورد ارزیابی واقع می‌شود خیلی از مسائل باید از منظر مشتری بررسی شود مانند میزان رضایتمندی مشتری، مشتریان سود ده، مشتریان کلیدی، تحویل کالا با سرعت بالا به مشتری، سرویس‌دهی‌های مناسب به مشتری همه‌ی این موارد باعث می‌شوند که از منظر مشتری به سازمان توجه شود.

۳- **منظر فرآیندهای داخلی (Internal Process):** کمبود مواد اولیه- سفارشات مانده در صف تولید- مشکلات کیفیتی- ساعات قطعی برق در سازمان، بهینه‌سازی فناوری‌های داخل شرکت و هرچیزی که وضعیت داخلی شرکت را بهبود ببخشد در این منظر قرار می‌گیرد مثلاً ارتباط مؤثر با سهامداران، تحویل به موقع کالا و .. درآمد بالقوه در بخش فروش.

۴- **منظر رشد و یادگیری (Leraning and Growth):** در این منظر از زوایه دید آموزش به شرکت نگاه می‌کنند افزایش مهارت کارکنان با آموزش‌های خاص و انطباق‌پذیری و سازگاری آنها در این منظر بررسی می‌شود. (کلاس‌های آموزشی جهت سخت افزار و نرم‌افزار)

۵- **منظر محیط جامعه (Environment/Community):** در این منظر حمایت از بیزینس یا کسب‌وکارهای محلی در آن بررسی می‌شود مثلاً حمایت از کارکنان آینده یعنی کسانی که بعداً ممکن است برای ما کار کنند در این منظر به آنها پرداخته می‌شود. (این منظر در بسیاری از کتاب‌ها وجود ندارد)

۶- **منظر رضایت کارکنان (Employee Satisfaction):** ایجاد فرهنگ سازمان سالم در این منظر بررسی می‌شود. نگهداری و حمایت از کارکنان کلیدی در این منظر بررسی می‌شود و اینجا سعی بر آن است که کارکنان کلیدی پروژه نگهداری شوند مثلاً افزایش پاداش ، وقتی در یک شرکت افزایش پاداش مطرح می‌شود در واقع از منظر رضایت کارکنان عملکرد سازمان سنجیده می‌شود.

(این منظر در بسیاری از کتاب‌ها وجود ندارد)

مثال: تعدادی از سنجه‌های یک سازمان را مثال بزنید و بگویید از کدام منظر می‌توان آن را مورد بررسی قرار داد؟

- زمان متوسط رسیدگی به شکایات، در سازمان زمان متوسط به شکایات بررسی می‌شود یعنی یک مدیر می‌تواند در سازمان خودش این موضوع را بررسی نماید. ما از منظر رضایت مشتری به موضوع نگاه می‌کنیم.

- درآمد هفتگی حاصل از فروش ، از منظر مالی است.

- تعداد درخواست های استخدام در شرکت ، از منظر کارکنان است.

- درصد کارکنانی که دارای ساعات کاری متغیر هستند ، از زاویه دید رضایت کارکنان می‌باشد.

از منظر محیط، میزان انتشار خروجی‌های محصول ما به محیط بیرونی چقدر است و یا میزان آب مصرفی به ازای هر واحد تولید یا هر کارمند چقدر است یا درصد پسماندهای ایجاد شده.

توجه: این منظرها خط جدایی از هم ندارند و گاهی با هم در ارتباط اند مثلاً وقتی درصد مشتریان بررسی می‌شود علاوه بر منظر مشتریان به نحوی به منظر مالی هم می‌پردازد چون وقتی مشتریان زیاد می‌شوند درآمد مالی سازمان هم بالا می‌رود. مثلاً تعداد مشتریانی که در شروع پروژه پرداخت نقدی داشته‌اند.

نکته: فراوانی سنجش از زاویه دید هر کدام از وجوه ، متفاوت است.

یعنی سنجه‌های موجود برخی ماهیانه سالیانه و هفتگی و برخی ساعتی اندازه‌گیری می‌شود. مثلاً درآمد هفتگی حاصل از فروش یا مثلاً جلب مشتری ممکن است بصورت فصلی یا ماهیانه فراوانی این سنجش در نظر گرفته شود.

مثال از منظر رضایت مشتری: میزان تماس‌های پاسخ داده شده یا نشده در اولین تماس بررسی شود تا میزان رضایت مشتریان تعیین شود. این سنجه می‌تواند روزانه ، ساعتی و لحظه‌ای بررسی شود. یا مدیر بیمارستان با کمک یک داشبورد وضعیت اورژانس خودش را به صورت لحظه‌ای می‌بیند (تعداد مراجعین - تصادفی ها - تعداد پزشکان - کارکنان) (این از منظر رضایت مشتری است اما لحظه‌ای باید اندازه‌گیری شود)

مثال از منظر رشد و آموزش: مهارت ICDL برای چند درصد کارکنان لازم است؟ چند درصد کارکنان ما زن هستند؟ چند درصد از کارکنان شامل رشد یادگیری شدند و مهارت‌هایشان افزوده شده و گواهینامه خام گرفتند. تعداد کلاس‌های آموزشی که در یکسال گذشته برگزار شده. میزان حضور کارکنان در کلاس های آموزشی. چند روز آموزش در سال برگزار شده است.

نکته: برخی سنجه‌ها در بعضی سازمان‌ها با اهمیت‌تر هستند مثلاً در دانشگاه‌ها منظر آموزش بسیار با اهمیت است مؤسسات مالی یا بانک‌ها از منظر مالی به بررسی سازمان خود می‌پردازند.

چشم‌انداز و فرآیند (Perspective And Progress):

Our perspectives and progress

FINANCIAL ☺ Utilization of assets ☹ Optimization of working capital, EBIT, growth, etc.	CUSTOMER FOCUS ☺ Increasing customer satisfaction ☹ Gaining profitable customers, etc.	ENVIRONMENT/ COMMUNITY ☺ Supporting local businesses ☹ Linking with future employees ☹ Community leadership, etc.
INTERNAL ☺ Delivery in full on time ☺ Optimizing technology ☹ Work accidents, etc.	EMPLOYEE SATISFACTION ☺ Positive company culture ☹ Retention of key staff ☹ Increased staff recognition, etc.	LEARNING AND GROWTH ☺ Increasing empowerment ☹ Increasing staff adaptability ☹ Coaching increasing, etc.

نمودار زیر یک گزارشی است که شش منظر کارت‌های متوازن را به کمک سنجه‌هایی که برای هر منظر تعیین کرده است بررسی می‌کند. سنجه‌ها دسته بندی و با شکلک‌ها نمایش داده شده و بررسی می‌شود:

مثلاً در Customer focus (تمرکز بر رضایت مشتری) ، گزینه اول یعنی افزایش رضایت مشتری با شکل ☺ نشان داده شده

است یعنی نه خوشحال است و نه ناراحت و این نشان می‌دهد این سنجه در شرکت ما در حد متوسط است. مثلاً منظر محیط / جامعه سه سنجه دارد که سنجه وسطی خوب نیست و با علامت ☹ نشان داده شده است.

نکته: BSC کارت‌های امتیازی متوازن مؤثرترین و کارآمدترین تکنیک حال حاضر دنیا می باشد یعنی در برنامه‌های راهبردی و استراتژی‌ها و ارزیابی جامعه عملکردشان با استفاده BSC بسیار موفق بوده است. اما شرکت‌های کوچک خیلی به BSC بستگی ندارند و در شرکت‌های ایرانی به دلیل عدم توجه به رقابت این سنجه‌ها کمتر وجود دارد.

عوامل کلیدی (یا حیاتی) موفقیت (CSF) Critical Success Factors: فاکتورهای بحرانی موفقیت ، آریه‌ای شفاف از سطوح کلیدی عملکرد سازمان است. لذا، فاکتورهای بحرانی موفقیت فعالیت‌های بادوامی را تعریف می‌کنند که سازمان باید در طول زمان به‌خوبی انجام دهد تا در راه رسیدن به مأموریت موفق شود.

این عوامل از منظرهای ۶ گانه عوامل کلیدی موفقیت متفاوت است:

مثلاً عوامل موفقیت کلیدی از منظر مالی: کاهش هزینه‌ها ، بهبود و افزایش جریان نقدینگی ، کسب یک رتبه عرضه کننده منتخب توسط مشتریان کلیدی ، معرفی و آریه خدمات جدید. مثلاً افزایش رضایت کارکنان یک عامل کلیدی موفقیت در منظر رضایت کارکنان. و اختصاص پاداش‌های مناسب ، محیط آموزش‌های مناسب ، برگزاری جشن در موفقیت‌هایی که کارکنان بدست می‌آورند ، ایجاد یک محیط کاری جالب و دلپذیر کاری برای کارکنان عامل کلیدی موفقیت هستند که از منظر رضایت کارکنان بررسی می‌شوند.

دسترسی آزادانه کارکنان به مسایل استراتژیک از منظر رشد و یادگیری است و یا امنیت و سلامت کار از منظر فرآیند های داخلی بررسی می‌شود و احساس مثبت عمومی در سازمان از منظر رضایت از محیط است.

شاخص‌های ارزشیابی سازمان

در هر شرکت یا سازمان برای سنجیدن نحوه عملکرد سیستم ، نیاز به یک سیستم ارزشیابی است. برای ارزشیابی باید شاخص (Indicator) تعیین کرد. به طور کلی شاخص‌ها به چهار دسته تقسیم می‌شوند:

۱- شاخص‌های کلیدی نتیجه (KRI) Key Result Indicator

شاخص‌هایی هستند که تعیین می‌کنند که یک عمل یا فعالیت چگونه انجام شده است. (از نظر زمانی مربوط به گذشته است)

هدف: نشان دادن نتیجه چندین عمل - حرکت در مسیر صحیح

مثال‌هایی از شاخص‌های کلیدی نتیجه:

- رضایت مشتری
- سود خالص قبل از کسر مالیات
- سودآوری مشتریان
- رضایت کارکنان
- بازده سرمایه مورد استفاده

شاخص‌های کلیدی نتیجه نشان می‌دهند که مثلاً اگر چند عمل فوق با هم انجام شوند چه نتیجه‌ای دارند و آیا سازمان در مسیر درست حرکت می‌کند یا خیر؟

۲- شاخص‌های نتیجه (RI) Result Indicator

شاخص‌هایی هستند که مشخص می‌کنند که چه فعالیتی انجام شده است. (از نظر زمانی مربوط به گذشته است)

معمولاً تمام سنج‌های مالی از این شاخص‌ها هستند.

به عنوان مثال:

- سود خالص یا راندمان
- برای تخت‌های بیمارستان با واحد هفته (مقدار زمانی که تخت خالی یا پر بوده است)
- رسیدگی به شکایات مشتریان کلیدی
- در مسافرت هفته گذشته چقدر بنزین مصرف شده است؟

۳- شاخص‌های کارایی (عملکرد) (PI) Performance Indicator :

شاخص‌هایی هستند که تعیین می‌کنند که هم‌اکنون و در حال حاضر چه فعالیت یا عملیاتی در حال انجام شدن است. (از نظر زمانی مربوط به زمان حال است)

نکته: - شاخص‌های عملکرد کلیدی نیستند

- کمک به هم راستا کردن تیم‌ها با استراتژی سازمان

- نمایش در کارت‌های امتیاز همراه با KPIs

شاخص‌های عملکرد کارشان این است که تیم‌ها را در راستای استراتژی که تصمیم گرفته شده است قرار می‌دهند پس استراتژی سازمان باید مشخص شود و با توجه به آن شاخص‌های کلیدی عملکرد هم می‌تواند تعیین شود چون می‌توان با کمک استراتژی به شاخص‌های عملکرد رسید
شاخص‌های عملکرد معمولاً در کارت‌های امتیاز متوازن همراه با KPI ها نشان داده می‌شود.

• مثلاً ۱۰ درصد از بهترین مشتریان ، چند درصد افزایش فروش برایمان داشته اند؟

• مثلاً شکایات مشتریان کلیدی

• مثلاً در حال حاضر در باک ماشین چه مقداری بنزین وجود دارد

۴- شاخص‌های کلیدی عملکرد (KPI) Key Performance Indicator :

شاخص‌هایی هستند که تعیین می‌کنند که چه عمل یا فعالیتی باید انجام شود تا کارایی بیشتر شود. این شاخص‌ها برای رقابت بسیار مهم هستند یعنی مشخص می‌کنند چه کارهایی باید انجام شود تا سازمان یا شرکت بتواند در یک رقابت ، خوب و قدرتمند ظاهر شود. (از نظر زمانی مربوط به آینده است)

- برای حرکت با اتومبیل به سمت مقصد مشخصی چه مقدار بنزین مورد نیاز است.

- تعداد بیماران در بخش اورژانس یک شاخص کلیدی عملکرد است. لذا توسط مدیر اورژانس به صورت لحظه‌ای بررسی می‌شود و در صورت افزایش بیش از اندازه باید تصمیمات لازم را به موقع اتخاذ نماید.

شاخص‌های کلیدی عملکرد دارای هفت خصوصیت (پارامتر) هستند. این هفت ویژگی عبارتند از:

۱- این شاخص‌ها سنجه‌هایی هستند که مالی نیستند غیر مالی هستند، یعنی به وسیله واحد پول بیان نمی‌شود پس سنجه‌های مالی سنجه‌های شاخص‌های کلیدی عملکرد نیستند. این امر در مدیریت بسیار حائز اهمیت می‌باشد. دلیل آن هم این است وقتی سنجه به شاخص‌های مالی مربوط می‌شود یعنی مربوط به گذشته است یعنی کارهای انجام شده و در تعریف شاخص کلیدی عملکرد می‌خواهیم کاری کنیم تا در آینده به عملکرد بهتری برسیم.

۲- به صورت مکرر ۷ روز هفته و ۲۴ ساعته اندازه‌گیری می‌شود، و اگر شاخصی مثلاً به صورت ماهانه باشد جزو شاخص‌های کلیدی عملکرد نخواهد بود. میزان سرمایه‌گذاری در تحقیقات فصلی است و شاخصی است که نمی‌تواند به عنوان شاخص کلیدی عملکرد باشد.

۳- CEO (مدیر ارشد اجرایی)، در واقع این سنجه‌ها حتماً باید زیر نظر مستقیم مدیر ارشد اجرایی بررسی و اندازه‌گیری شوند و پشتیبانی مدیر ارشد اجرایی را لازم دارند.

۴- توسط تمام کارکنان باید شناخته باشد، و شیوه‌های اقدامات اصلاحی آن باید مشخص باشد (بطور واضح و روشن)

۵- مسئولیت‌های افراد و تیم‌های کاری باید کاملاً مشخص شود، در شاخص‌های کلیدی عملکرد بسیار مهم است.

۶- باید اثری مهم و معنی‌دار داشته باشند، یعنی حتماً باید در عوامل کلیدی موفقیت CSF تأثیر بگذارند پس حتماً باید دارای اثر قابل توجهی باشد و حتماً چیزی فراتر از یک منظر تعریف شده در یک کارت امتیازی متوازن باشد. مثلاً سود حاصل از افزایش مشتری هم در بخش رضایت مشتری است و هم در بخش مالی.

۷- دارای اثرات مثبتی باشند، شاخص‌های کلیدی عملکرد روی تمام سنجه‌های عملکرد دیگر اثر مثبت دارند به عبارت دیگر وقتی شاخص‌های کلیدی عملکرد بهبود می‌بخشیم روی رفتار بقیه عملکردها تأثیر مثبت می‌گذارد. پس یک شاخص کلیدی عملکرد خوب باید با سایر سنجه‌ها ارتباط داشته باشد مثلاً رضایت مشتری می‌تواند در حوزه‌های مختلفی از سازمان اثری مثبت داشته باشد.

نکته: شاخص‌های کلیدی عملکرد معمولاً متعلق به زمان حال هستند و در مقابل سنجه‌های گذشته قرار می‌گیرند و این سنجه‌ها به ما می‌گویند نحوه عملکرد ما در آینده چگونه می‌تواند باشد.

سؤال: آیا رضایت مشتری می‌تواند در سازمانی شاخص کلیدی نتیجه باشد و در سازمان دیگری خیر؟ و آیا می‌توانید شاخصی را مثال بزنید که در هر سازمانی یک شاخص مشخص باشد و نتوان جابجایش کرد؟

نکته: در یک سازمان یک شاخص می‌تواند جزو شاخص‌های نتیجه باشد اما در سازمانی دیگر جزو شاخص‌های عملکرد قرار گیرد چون سازمان‌ها استراتژی‌ها و چشم‌اندازهای متفاوتی دارند پس شاخص‌های متفاوتی هم خواهند داشت.

قانون ۱۰/۸۰/۱۰

این قانون بیان می کند که از ۴ نوع شاخص موجود ، از بین شاخص هایی که در یک سازمان تعیین می شوند ۸۰٪ مربوط به شاخص های نتیجه و عملکرد ، ۱۰٪ آن ها مربوط به شاخص های کلیدی نتیجه و ۱۰٪ شاخص های کلیدی عملکرد هستند. یعنی اگر در یک شرکت ۵۰ شاخص تعیین شود طبق این قانون از این شاخص ها تنها ۵ مورد از شاخص ها شاخص کلیدی عملکرد می باشند.

Mission / Vision / Strategy

مأموریت (Mission): معمولاً مأموریت مانند یک فانوس دریایی یا یک راهنما است چون وقتی مأموریت مشخص می شود ممکن است ما هرگز به نتیجه بیان شده در مأموریت نرسیم یعنی یک هدف غایی است. یعنی مثلاً هر کاری که در جهت این مأموریت انجام شود به نحوی ما را به سوی هدف غایی ما سوق می دهد و به گونه ای شعار گونه است اما به ما انگیزه می دهد. ممکن است یک مأموریت مدت ها باز بماند چون یک مسئله کلی است.

مثلاً مأموریت ما: تأمین انرژی لازم در زمان صحیح با قیمت صحیح مثلاً در شرکت توزیع برق ممکن است این به صورت یک مأموریت تعریف شود.

چشم انداز (Vision): نشان می دهد ما به کجا می رویم و سازمان چه می خواهد باشد و اگر سازمان موفق شد کاری را که در مأموریتش تعیین کرده انجام داد چه چیزی را کسب کرده است؟ پس چشم انداز می تواند سازمان را به هیجان وادارد و آن را تحریک کند (به شرطی که به قدر کافی روشن بیان شده باشد) چشم انداز نمی تواند مانند مأموریت شعار گونه باشد و کاملاً باید به روشنی بیان شود یعنی کمی باید از مأموریت روشن تر باشد.

چشم انداز محدود به زمان است و توسط مدیران ارشد سازمان حمایت می شود.

مثلاً چشم انداز ۵ سال آینده: تأمین کننده برتر انرژی در منطقه (باید محدوده و حوزه داشته باشد)

راهبرد (Strategy): مسیری که باید طی شود تا به چشم انداز رسید یعنی کاری که باید انجام داد تا به چشم انداز رسید.

مثلاً استراتژی سازمان ممکن است بر روی موارد زیر تمرکز داشته باشد:

(۱) جذب مشتریان سودآور (۲) افزایش کارایی و اثر بخشی هزینه

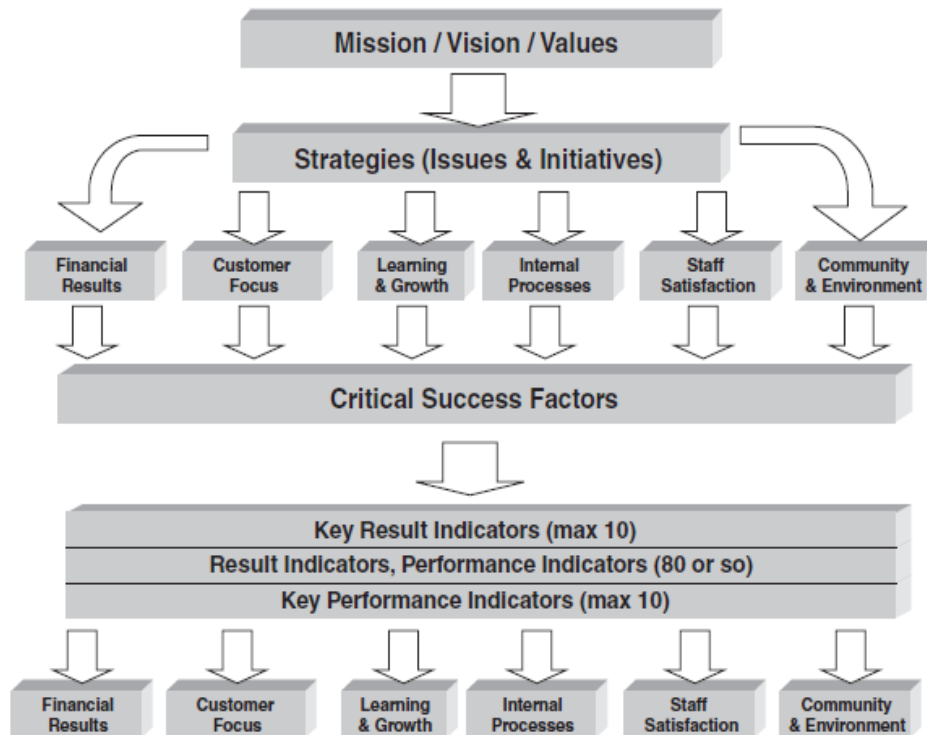
(۳) خلاقیت کارکنان (۴) استفاده از تجربیات موفق

یعنی ارتقاء یا کاهش این ۴ موضوع حتماً باید بررسی شود تا بتوان به چشم انداز دست یافت.

گاهی به سه واژه Value ، Mission/Vision/Strategy هم گفته می‌شود.

تشخیص این شاخص‌های کلیدی عملکرد کار ساده‌ای نیست اما افراد با تجربه بالا در ذهنشان شاخص‌های بسیار زیادی دارند بعضی از شاخص‌ها را می‌توان با تحلیل‌های هوشمند استخراج کرد. این موارد در شکل زیر کاملاً توضیح داده شد.

در بالای شکل مأموریت و چشم‌انداز هستند و بعد استراتژی‌ها تعریف می‌شوند و استراتژی‌ها در ۶ منظر بررسی می‌شوند. هر یک از این ۶ منظر عوامل کلیدی موفقیت را تحت تأثیر قرار می‌دهند و این عوامل در سنجه‌ها گنجانده می‌شوند که طبق قانون ۱۰/۸۰/۱۰ تفکیک می‌شوند.



داشبوردها

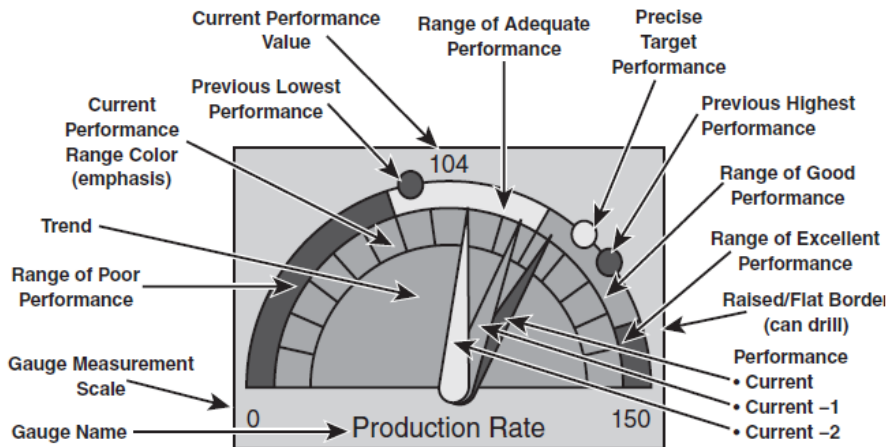
« داشبورد نموداری است که نشان دهنده‌ی مقادیر شاخص‌ها به شکل گرافیکی می‌باشد »

مثلاً بر روی داشبورد ماشین شاخص‌های مهمی وجود دارند که لازم است از آنها با خبر بود ، مانند مقدار بنزین خودرو یا سرعت.

این داشبوردها طراحی‌ها و رنگ‌های زیبایی دارند و در قالب نمودارهای متفاوتی بیان می‌شوند مثل نمودار های راداری. نموداری که انتخاب می‌شود نسبت به کاربرد آن باید نمودار مناسبی باشد.

داشبوردها معمولاً به شش یا نه قسمت تقسیم می‌شوند. داشبوردهای ۶ قسمتی بهتر هستند چون باید اطلاعات کامل و جامع تنها در قالب یک صفحه نمایش داده شوند بهتر است خلاصه‌تر و گویاتر باشند.

شکل زیر یک سرعت‌نماست که چقدر اطلاعات مفید و جالبی دارد و این سرعت‌نماها بهتر است در تمام داشبوردها وجود داشته باشند مثلاً یک علامت سرعت‌نمای اتومبیل- در شاخص‌ها تشخیص اینکه کدام شاخص مهمتر است برای قرارگیری روی داشبورد نکته بسیار مهم و حائز اهمیت می‌باشد.



Total Cash Collections Gauge



Patient AR % Gauge



Total AR Balance Gauge



در شکل زیر داده‌های مکان محور در قاب یک داشبورد نمایش داده شده این داشبوردها به شکل آنلاین و لحظه‌ای آپدیت می‌شوند.

