



# Business Intelligence

---

Skorkovský  
KAMI, ESF MU



# Principy BI

---

- zpracování velkých objemů dat tak, aby výsledek této akce manažerům pomohl k rozhodování při řízení procesů
- výsledkem zpracování musí být relevantní informace, kterou dostanou manažeři ve správném čase
- základní zdroj dat, která se často ukládají do datových skladů jsou ERP systémy (relační DB)
- získání informací jako výsledek strukturovaných dotazů musí probíhat rychle (krátká odezva)
- používá se pro řízení na strategické, taktické u operační úrovni



# Principy BI

---

- **Definice 1** : BI je sběr a analýza dat, jejímž cílem je lepší porozumění a reakce na změny, kterým organizace neustále čelí
- **Definice 2** : BI je znalost podniku získaná za použití HW a SW technologií, která umožní přeměnit data organizace v informaci
- **Definice 3** : sada procesů, aplikací a technologií, jejichž cílem je účinně a účelně podporovat rozhodovací procesy ve firmě. Tyto procesy podporují analytické a plánovací činnosti podniků a organizací a jsou postaveny na principech multidimenzionálních pohledů na podniková data



# Nástroje (terminologie) BI

---

- ERP systémy –zdroj primárních dat (transakce-položky v NAV)
- Dočasná úložiště (DSA: **D**ata **S**taging **A**rea)
- Operativní úložiště (ODS : **O**perational **D**ata **S**tore)
- Transformační nástroje (ETL : **E**xtraction **T**ransformation **L**oading)
- Integrovaní nástroje (EAI : Enterprise Application Integration)
- Datové sklady
- Datová tržiště
- OLAP= On-Line-Analytical-Processing
- Reportingové nástroje
- EIS (Executive Information System)
- Data Mining

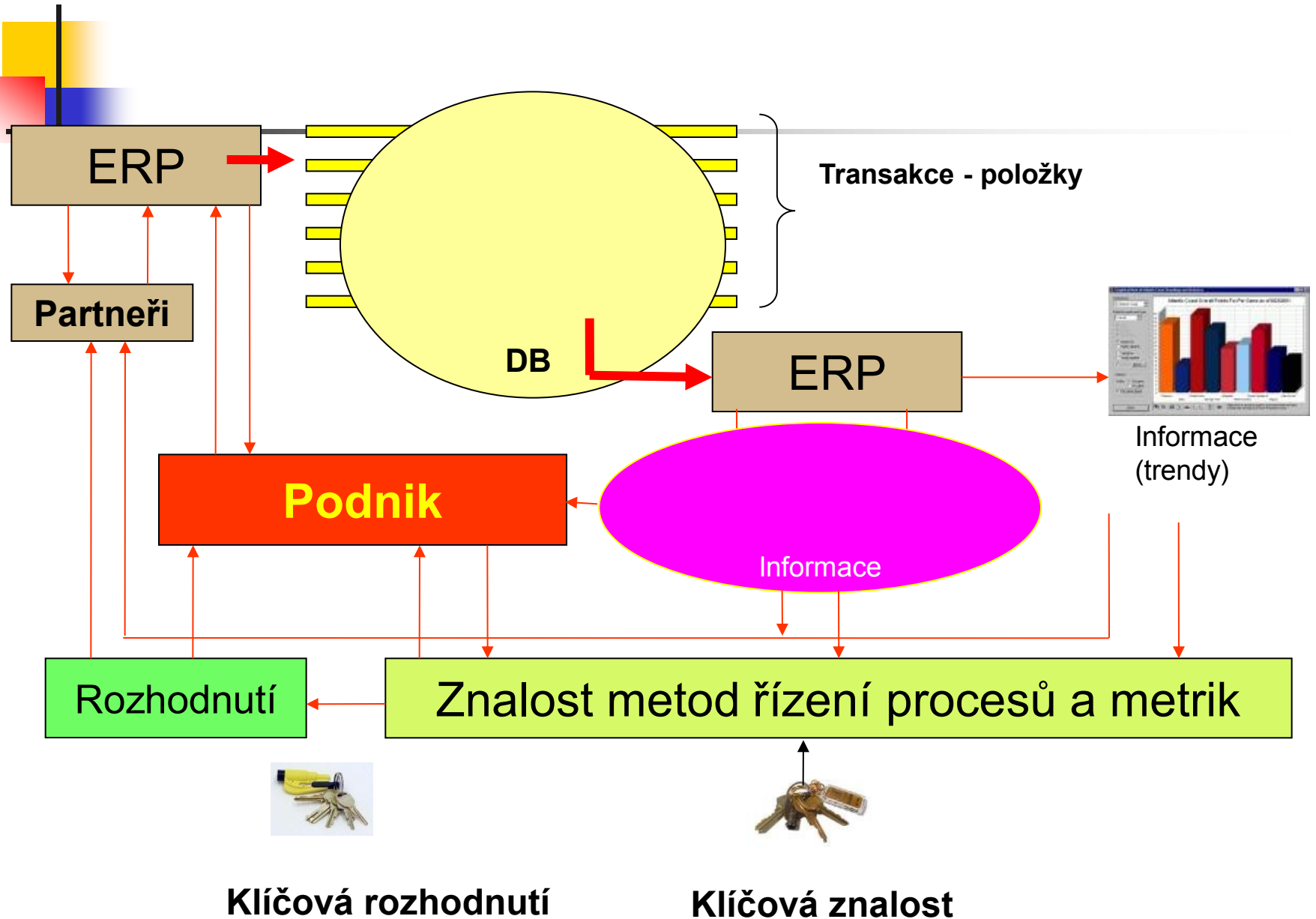


# Omezení ERP jako poskytovatele dat

---

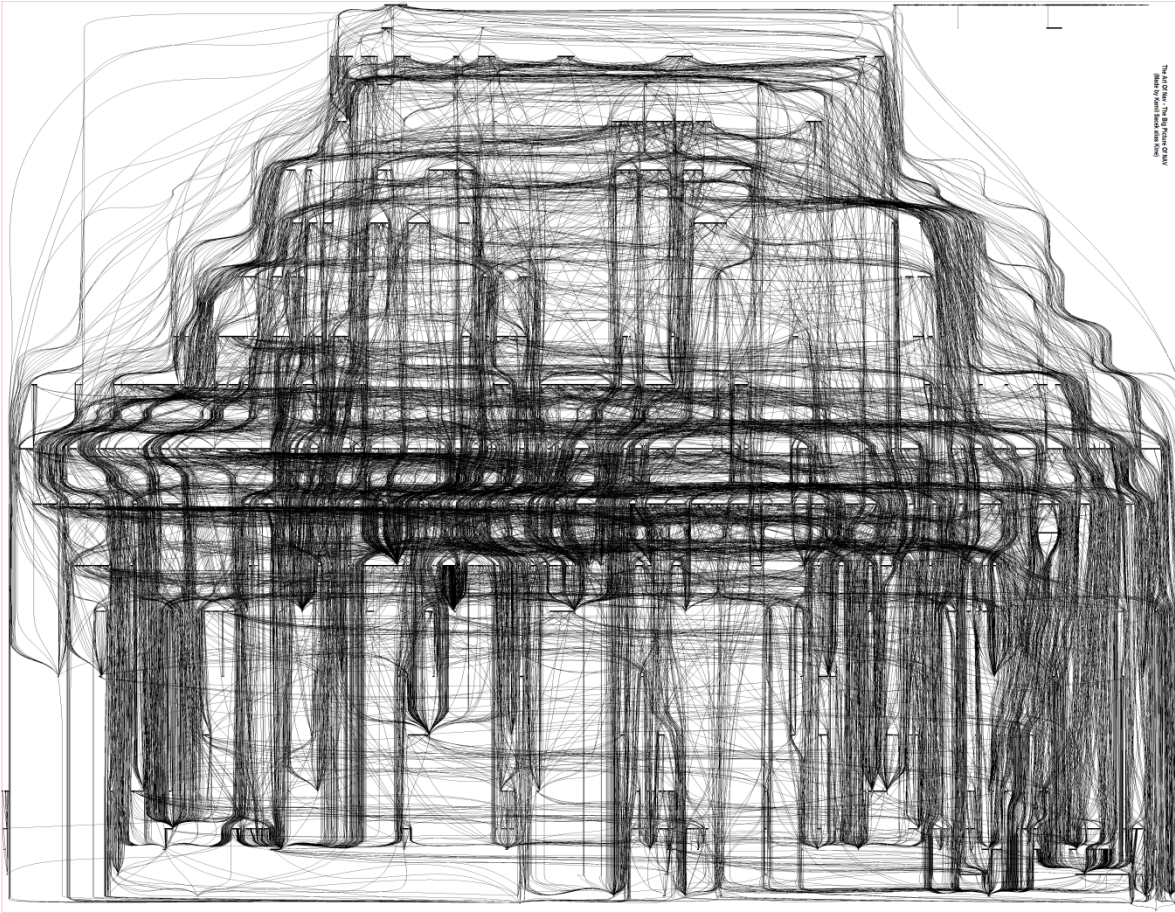
- Neumožňují rychle a pružně měnit kritéria výběru
- Není okamžitý přístup uživatelů k velkým objemům agregovaných dat
- ERP jsou primárně určeny k pořizování dat a jejich aktualizaci (relační typ databáze – pomalý přístup k požadovaným údajům)
- V každém podniku se objem dat za každých pět let zdvojnásobí, což ovšem také znamená, že systém je zahlcen redundantními daty
- Vícedimenzionální pohled na data v ERP je problematický. DB ERP není pro tento pohled stavěná. Databáze, které vzniknou přeměnou primárních dat z ERP a jsou využívány např. OLAP technologií jsou pro operace typu *drilling* optimalizovány

# Zjednodušené schéma využívání ERP





# Relations among tables I

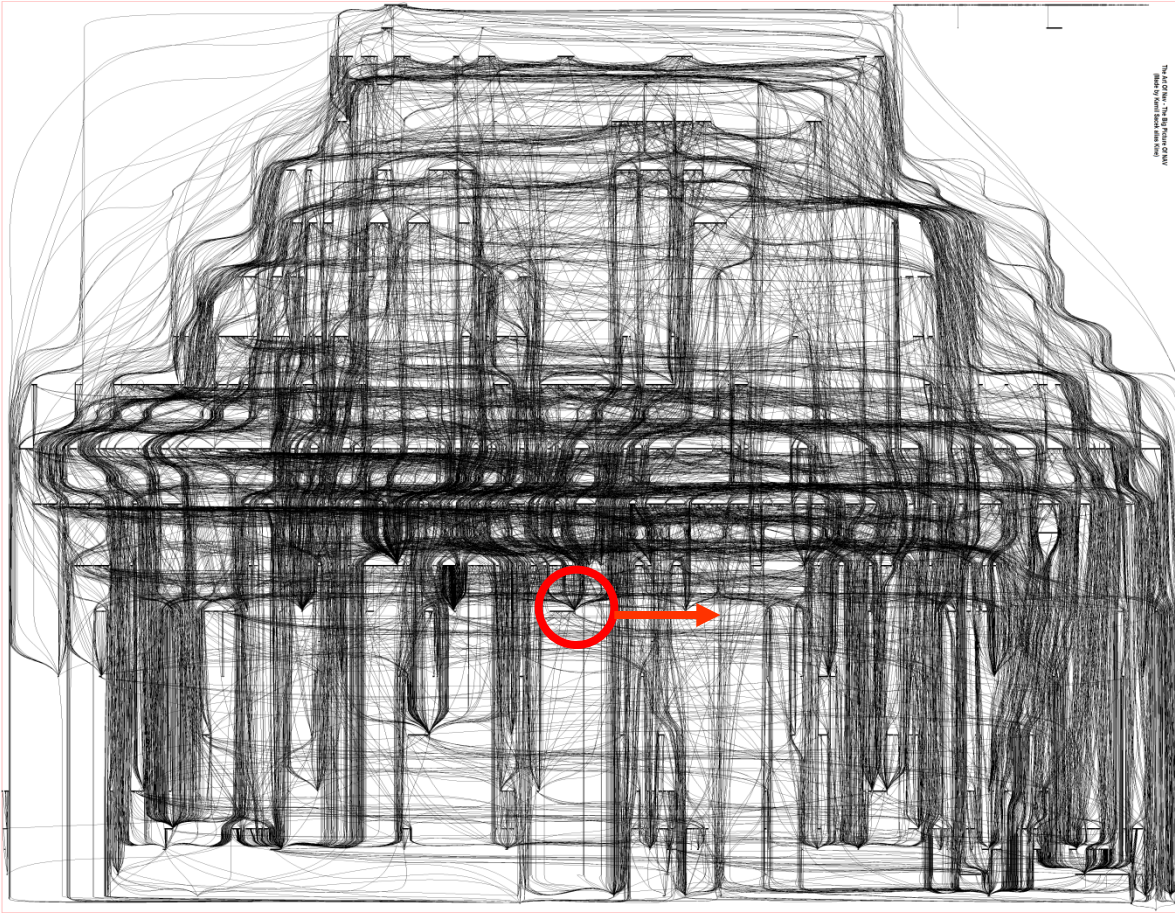


18.4.2018





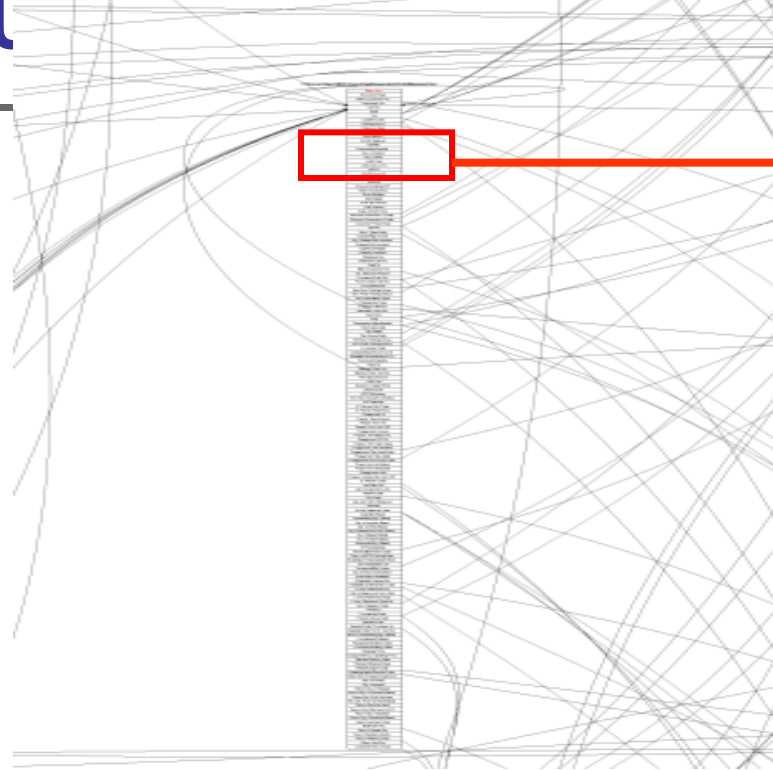
# Relations among tables II



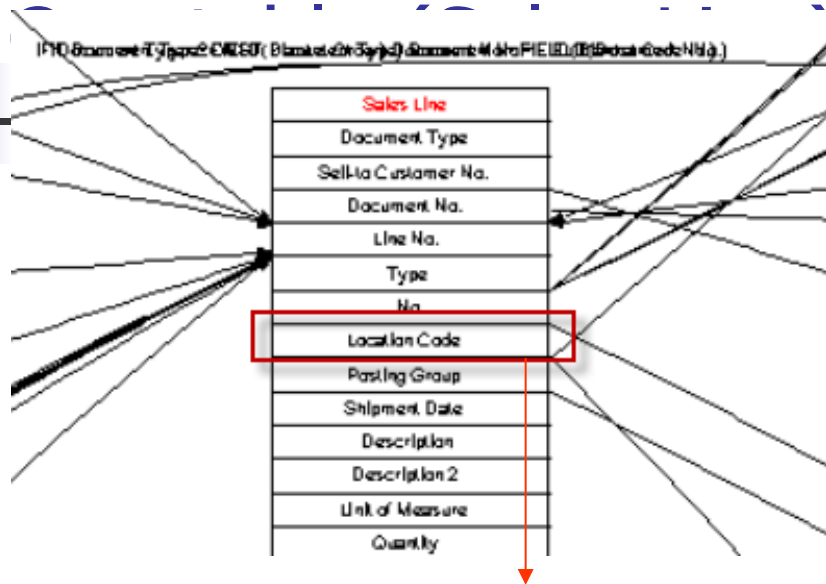
18.4.2018



# One table and its relations I



# and its relations II



ERP NAV

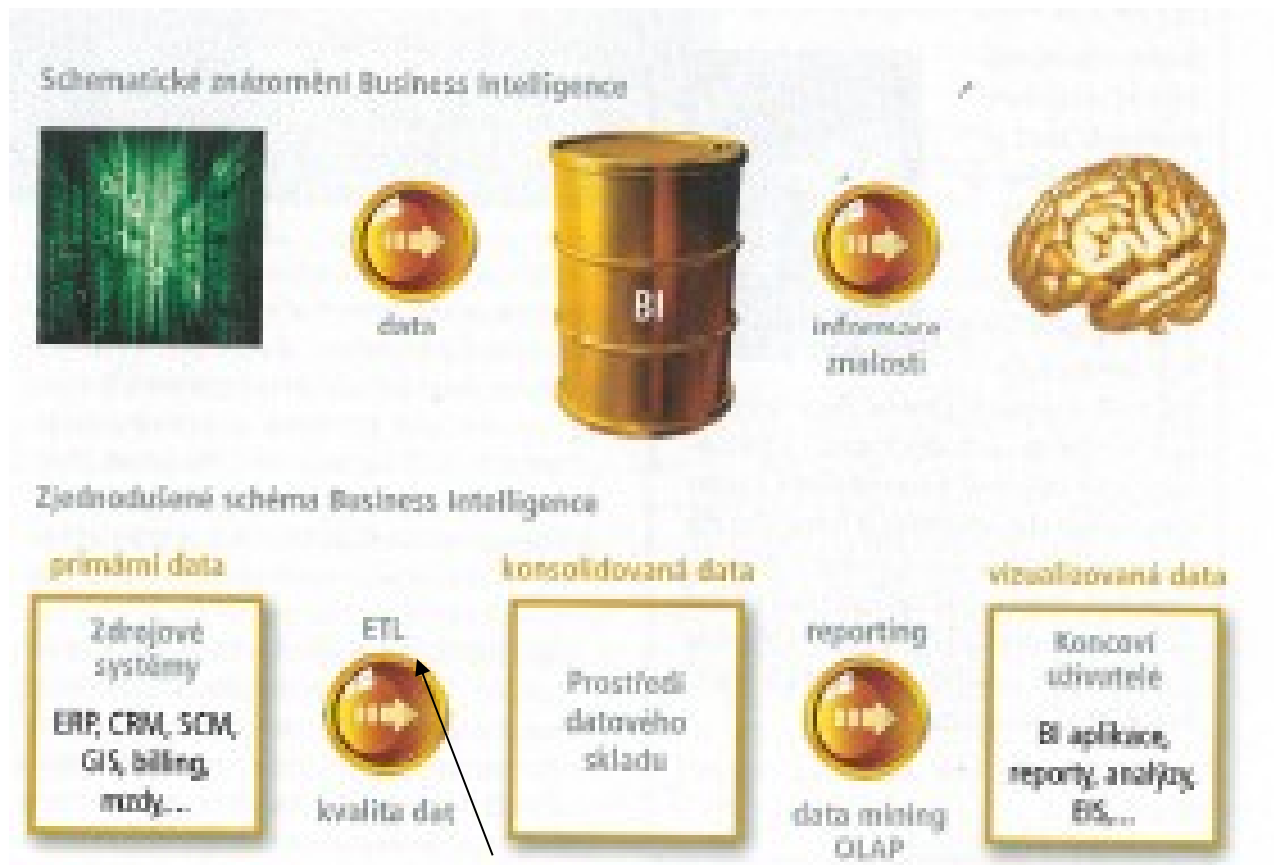
| Type | No.    | Purch... Code | Drop Shipment            | Description              | Location Code | Quantity | Reserved Quantity | L   |
|------|--------|---------------|--------------------------|--------------------------|---------------|----------|-------------------|-----|
| Item | 1980-S |               | <input type="checkbox"/> | MOSCOW Swivel Chair, red | ZELENÝ        | 6        |                   | k ^ |

Order Line Functions Posting Print Help

Item Information

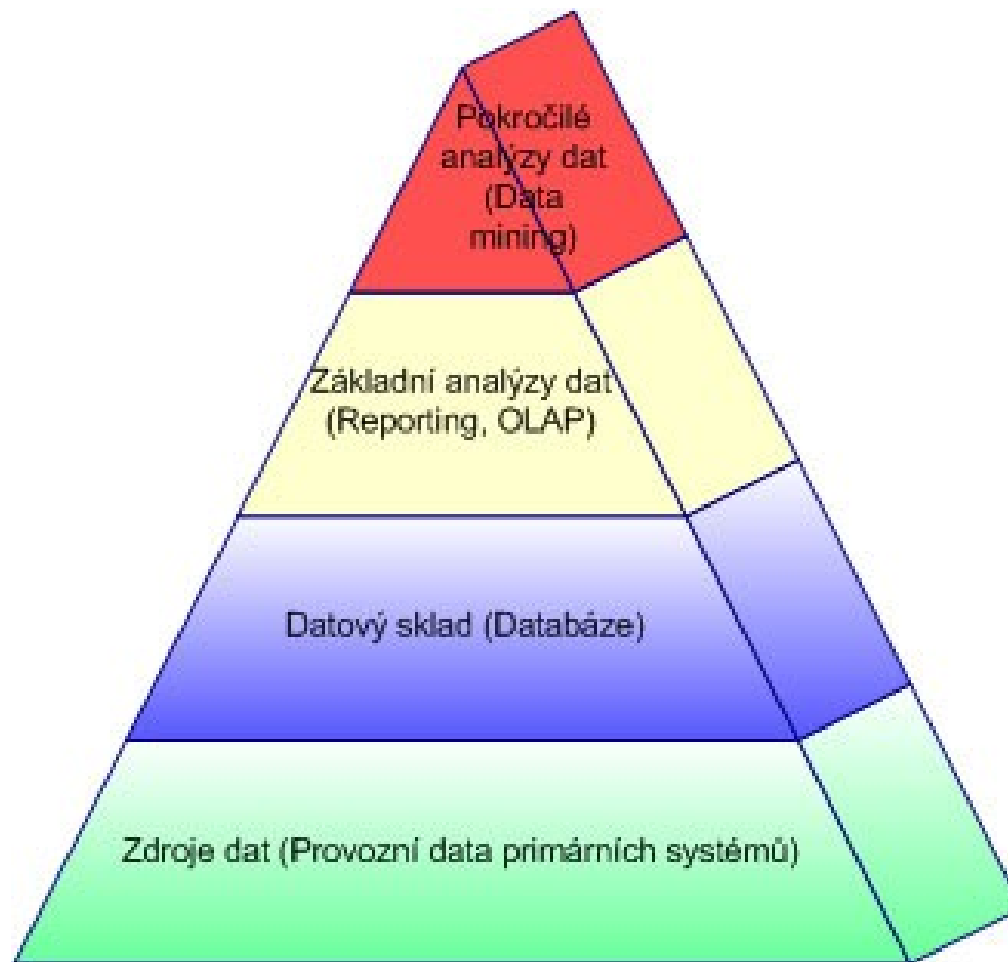
- Item Card
- Availability (8)
- Substitutions (1)
- Sales Prices (0)
- Sales Line Dis... (1)

# Schematické a zjednodušené schéma BI



Extraction Transformation Loading = datová pumpa

# Jiný pohled na BI



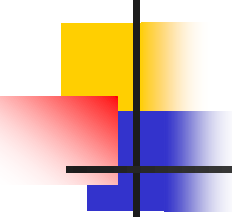


# Konstatování

---

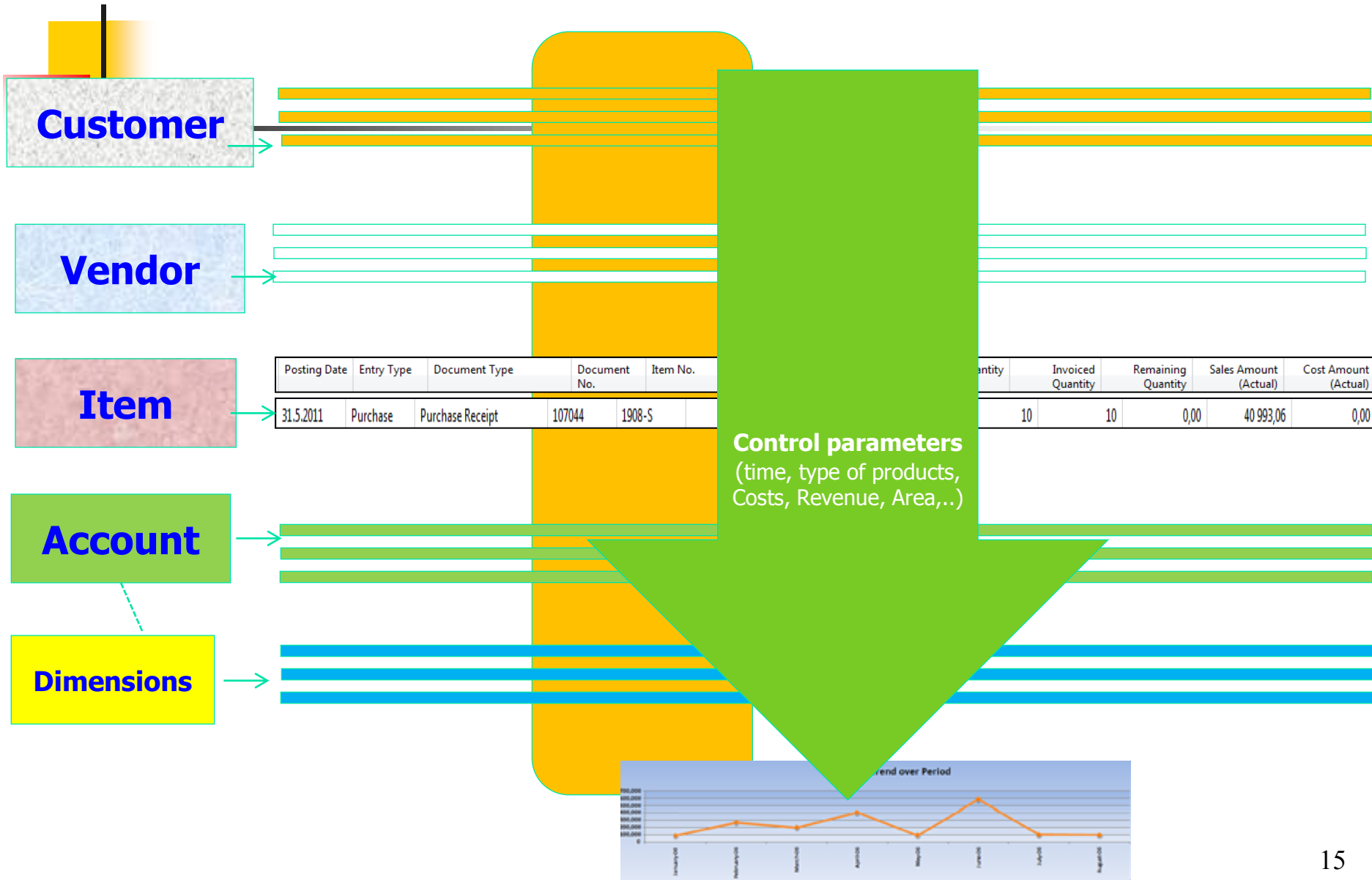
- V poslední dekádě minulého tisíciletí firmy hýčkaly mantru **ERP**. Důvodem byla utkvělá představa, že miliony utracené za licence a implementační služby se projeví tak, že pro uživatele už nebude chytrost žádné čáry a vše půjde jako po másle.... A to navždy.....

# Reporting (NAV tools or JETs)





# Main principles (source tables and their entries)



# Some chosen analysis – ERP MS Dynamics NAV 2009

## Working capital – setup of the accounting schedule from NAV

WC Working Capital Analysis - Účetní schéma

Název ..... WC

| Číslo řady | Popis                    | Typ s... | Součet    | Typ řady | T... Z... | Tučně | Nová strana |
|------------|--------------------------|----------|-----------|----------|-----------|-------|-------------|
|            | Working Capital Analysis | Účty     |           | Pohyb    | Č.. A..   |       |             |
|            |                          | Účty     |           | Pohyb    | Č.. A..   | ✓     |             |
| 1          | Goods all                | Součty   | 139995    | Pohyb    | Č.. A..   |       |             |
| 2          | Receivables              | Součty   | 315995    | Pohyb    | Č.. A..   |       |             |
| 3          | Cash total               | Součty   | 211995    | Pohyb    | Č.. A..   |       |             |
| 4          | Banks all                | Součty   | 221995    | Pohyb    | Č.. A..   |       |             |
| TR         | Receivables              | Vzorec   | 1+2+3+4+5 | Pohyb    | Č.. A..   | ✓     |             |
|            |                          | Účty     |           | Pohyb    | Č.. A..   |       |             |
| 5          | Short bank credits       | Účty     | 231000    | Pohyb    | Č.. A..   |       |             |
| 6          | Splatné závazky, celkem  | Součty   | 321995    | Pohyb    | Č.. A..   |       |             |
| 7          | Prepayments              | Účty     | 324100    | Pohyb    | Č.. A..   |       |             |
| 8          | Associate liabilities    | Součty   | 368995    | Pohyb    | Č.. A..   |       |             |
| TP         | Payables                 | Vzorec   | 5+6+7+8   | Pohyb    | Č.. A..   | ✓     |             |
|            |                          | Účty     |           | Pohyb    | Č.. A..   |       |             |
| ▶          | Working capital          | Vzorec   | TR TP     | Pohyb    | Č.. A..   | ✓     |             |

Účetní sc...    Funkce    Tisk    Nápověda

# Some chosen analysis

- Working capital – Show of the results from NAV

WC - WC - Náhled účetního schéma

Obecné Filtry dimenzí Možnosti

Název účetního schém... WC

Název rozložení sloupce . WC

Filtr data . . . . . 01.01.08..31.03.08

Filtr rozpočtu . . . . .

| Číslo řady | Popis                           | Net change (0)        | Net change (-3M)     | Net change (-6M)      | Net change (-9M)     |
|------------|---------------------------------|-----------------------|----------------------|-----------------------|----------------------|
|            | <b>Working Capital Analysis</b> |                       |                      |                       |                      |
| 1          | Goods all                       | 22 734 286,66         | 22 868 622,95        | 35 133 928,96         | 34 203 102,87        |
| 2          | Receivables                     | 8 896 962,33          | 15 620 590,10        | 13 095 293,10         | -2 424 454,90        |
| 3          | Cash total                      | -28 167,60            | 36 920,70            |                       |                      |
| 4          | Banks all                       | 7 249 513,07          | -161 688,80          |                       | -2 694 006,24        |
| <b>TR</b>  | <b>Receivables</b>              | <b>30 460 492,90</b>  | <b>57 101 259,84</b> | <b>34 005 343,10</b>  | <b>25 923 710,64</b> |
| 5          | Short bank credits              | -8 392 101,56         | 18 736 814,89        | -14 223 878,96        | -3 160 931,09        |
| 6          | Splatné závazky, celkem         | -12 766 412,03        | -9 171 504,29        | -3 876 790,60         | 295 083,40           |
| 7          | Prepayments                     |                       |                      |                       |                      |
| 8          | Associate liabilities           |                       |                      |                       |                      |
| <b>TP</b>  | <b>Payables</b>                 | <b>-21 158 513,59</b> | <b>9 565 310,60</b>  | <b>-18 100 669,56</b> | <b>-2 865 847,69</b> |
|            | <b>Working capital</b>          | <b>9 301 979,31</b>   | <b>66 666 570,44</b> | <b>15 904 673,54</b>  | <b>23 057 862,95</b> |

1 7 31 3 12

Funkce  Nápověda

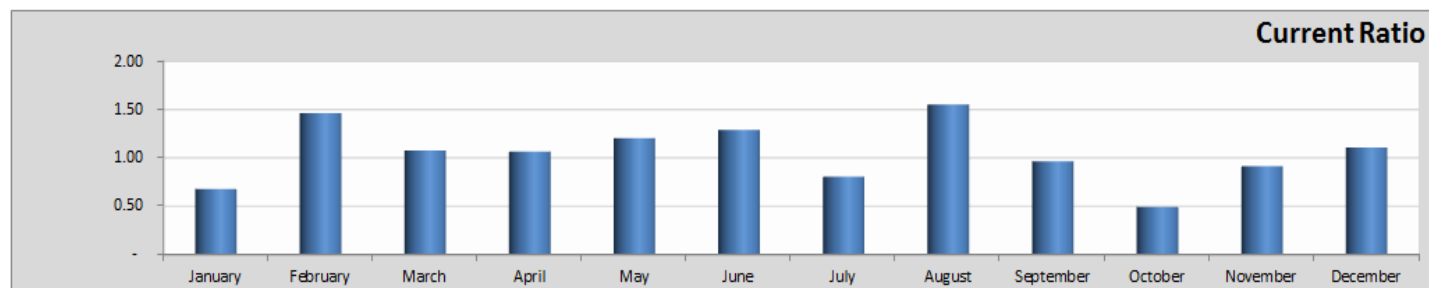
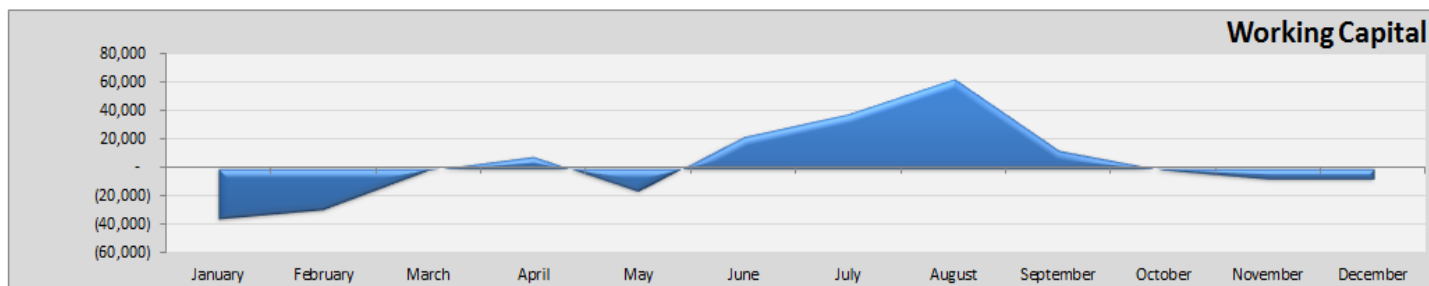
# Some chosen analysis

- Working capital – Show of the results from JETs=JET Reports

## Working Capital & Current Ratio

Report Date 8/23/2011

| 2011                |          |          |          |         |          |         |           |         |           |         |          |          |
|---------------------|----------|----------|----------|---------|----------|---------|-----------|---------|-----------|---------|----------|----------|
|                     | January  | February | March    | April   | May      | June    | July      | August  | September | October | November | December |
| Current Assets      | 74,405   | (90,939) | (16,110) | 136,096 | (92,528) | 101,144 | (155,777) | 174,615 | (363,170) | 1,015   | 72,525   | (72,789) |
| Current Liabilities | 109,902  | (62,118) | (14,989) | 127,587 | (76,890) | 78,566  | (193,757) | 112,467 | (376,168) | 2,070   | 79,494   | (65,841) |
| Working Capital     | (35,497) | (28,821) | (1,121)  | 8,508   | (15,638) | 22,579  | 37,980    | 62,148  | 12,998    | (1,055) | (6,969)  | (6,948)  |
| Current Ratio       | 0.68     | 1.46     | 1.07     | 1.07    | 1.20     | 1.29    | 0.80      | 1.55    | 0.97      | 0.49    | 0.91     | 1.11     |



# Some chosen analysis

## Inventory - Dashboard

**Company**

JetCorp, Belgium

JetCorp, North America

**Location**

Amsterdam Warehouse 1

Amsterdam Warehouse 2

**Item Category**

Awards & Recognition

Bags & Totes

**Product Group**

Flashlights

USB Drives

**Year**

2007 2008

2009 2010

**Quarter**

2010 Q1 2010 Q2

2010 Q3 2010 Q4

**Month**

2010 May 2010 Jun

2010 Jul 2010 Aug

**Inventory Value**

Millions

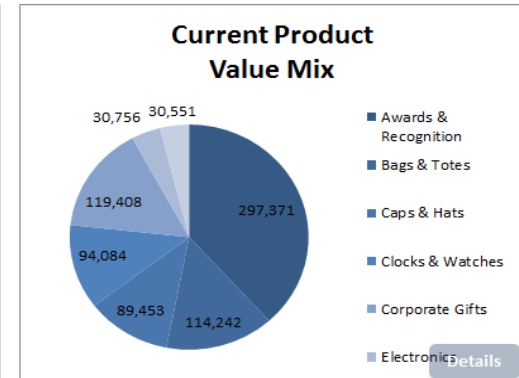
2010 Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

2010 Q1 2010 Q2 2010 Q3 2010 Q4

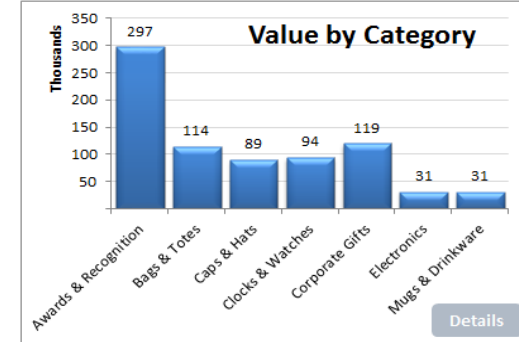
2010

- Awards & Recognition
- Bags & Totes
- Caps & Hats
- Clocks & Watches
- Corporate Gifts
- Electronics
- Mugs & Drinkware

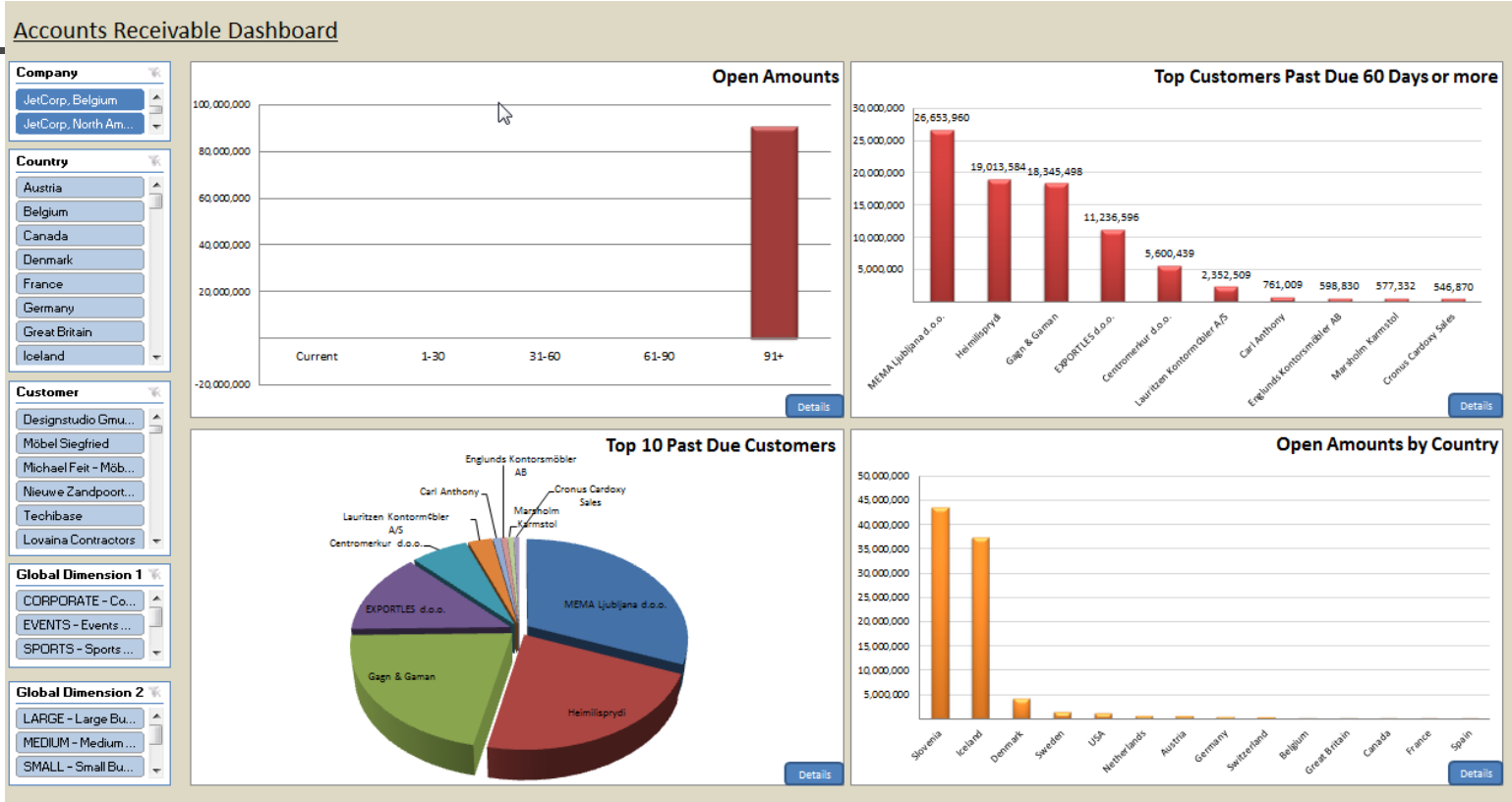
Details



| Average Inventory Value | Period  | Prev. Yr | Variance | %       | Grand Total |
|-------------------------|---------|----------|----------|---------|-------------|
| Awards & Recognition    | 188,840 | 200,797  | -11,957  | -5.95%  | 188,840     |
| Bags & Totes            | 72,551  | 83,459   | -10,908  | -13.07% | 72,551      |
| Caps & Hats             | 56,765  | 63,983   | -7,218   | -11.28% | 56,765      |
| Clocks & Watches        | 57,043  | 63,353   | -6,310   | -9.96%  | 57,043      |
| Corporate Gifts         | 73,856  | 81,520   | -7,664   | -9.40%  | 73,856      |
| Electronics             | 19,293  | 18,451   | 842      | 4.56%   | 19,293      |
| Mugs & Drinkware        | 19,052  | 16,732   | 2,320    | 13.86%  | 19,052      |

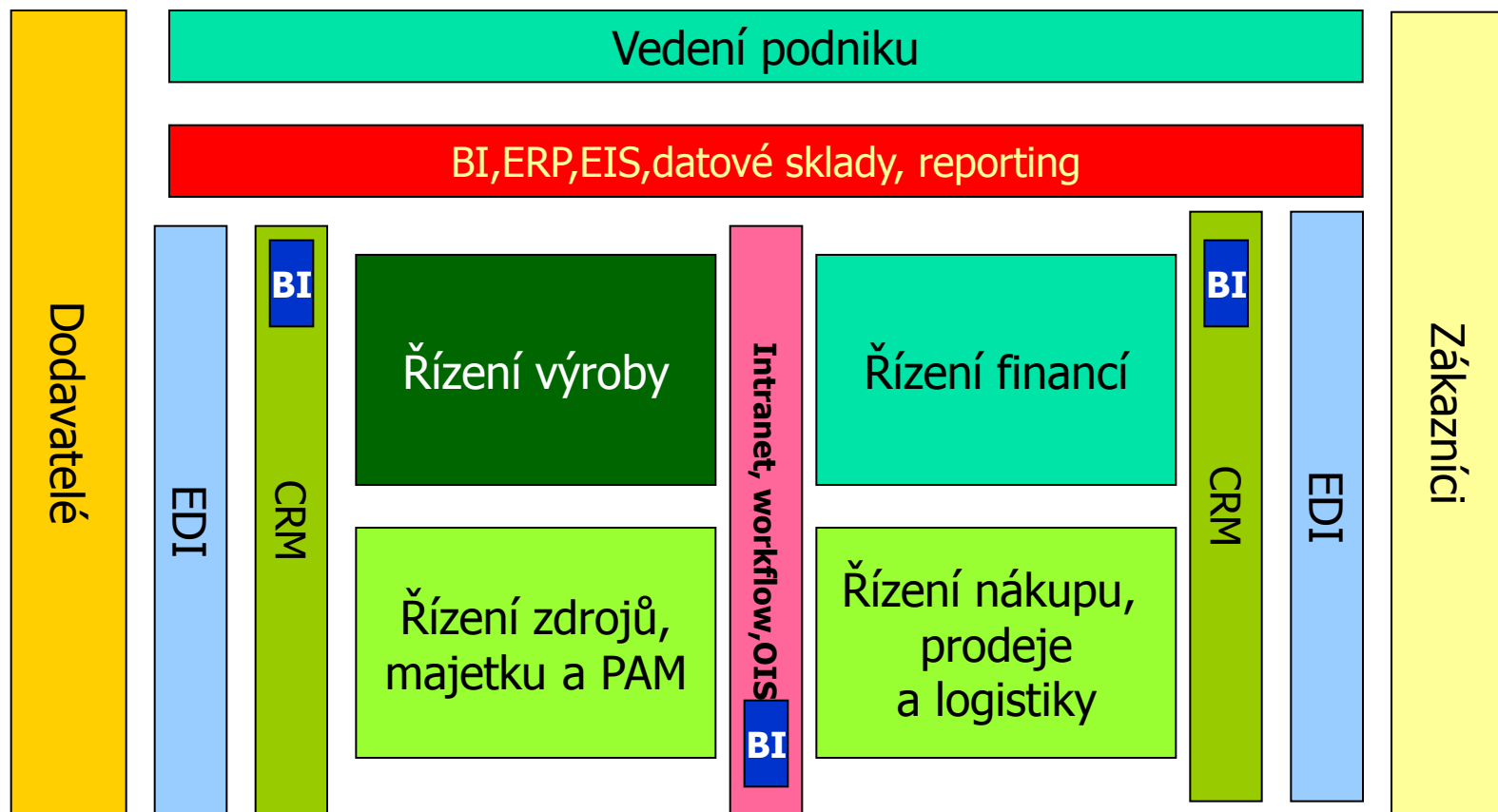


# Some chosen analysis examples (JETs)



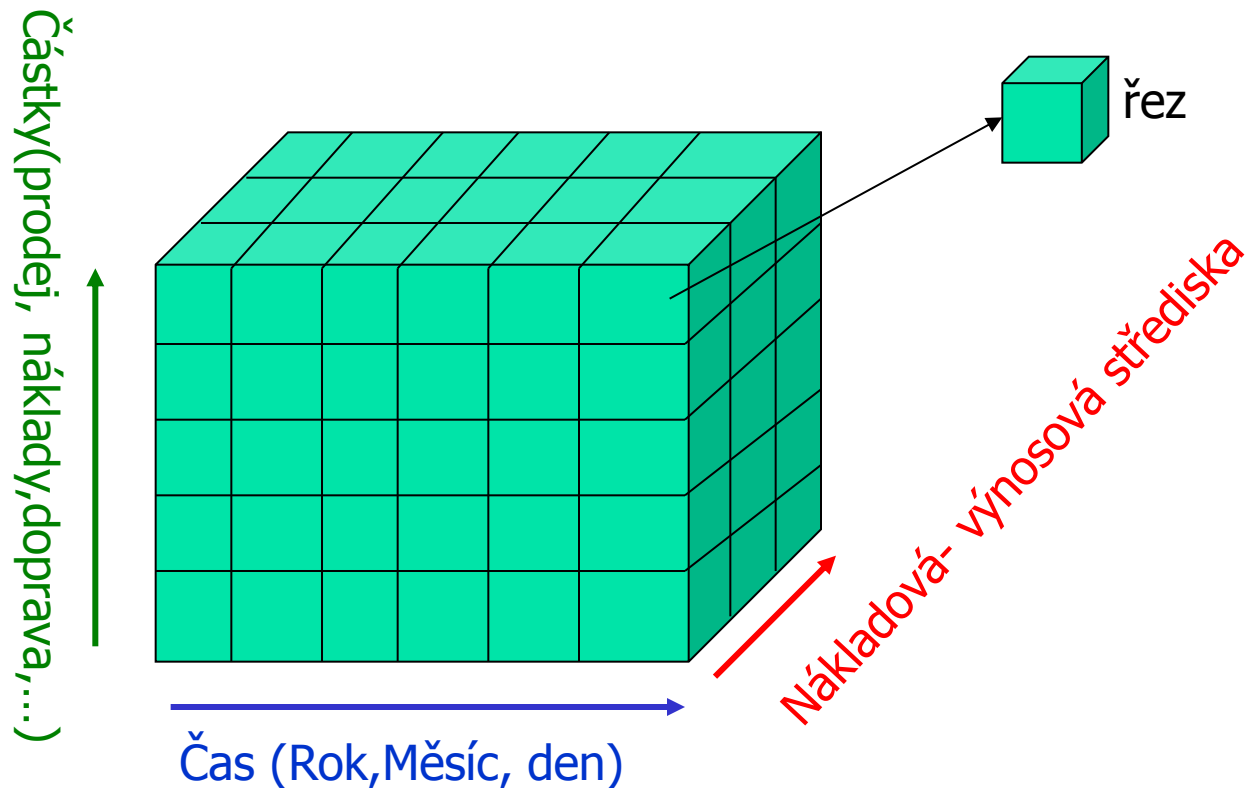


# Organizačně-technologické schéma podniku

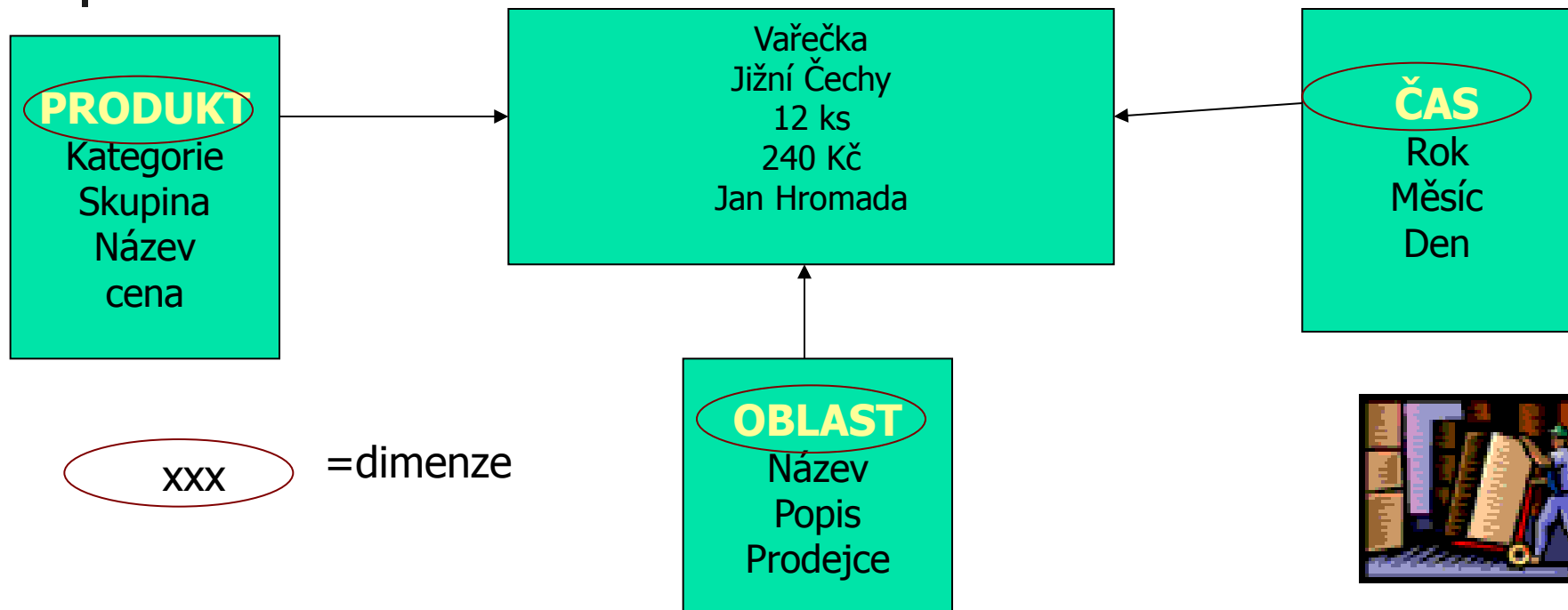


# OLAP kostka-adhoc definice dotazů

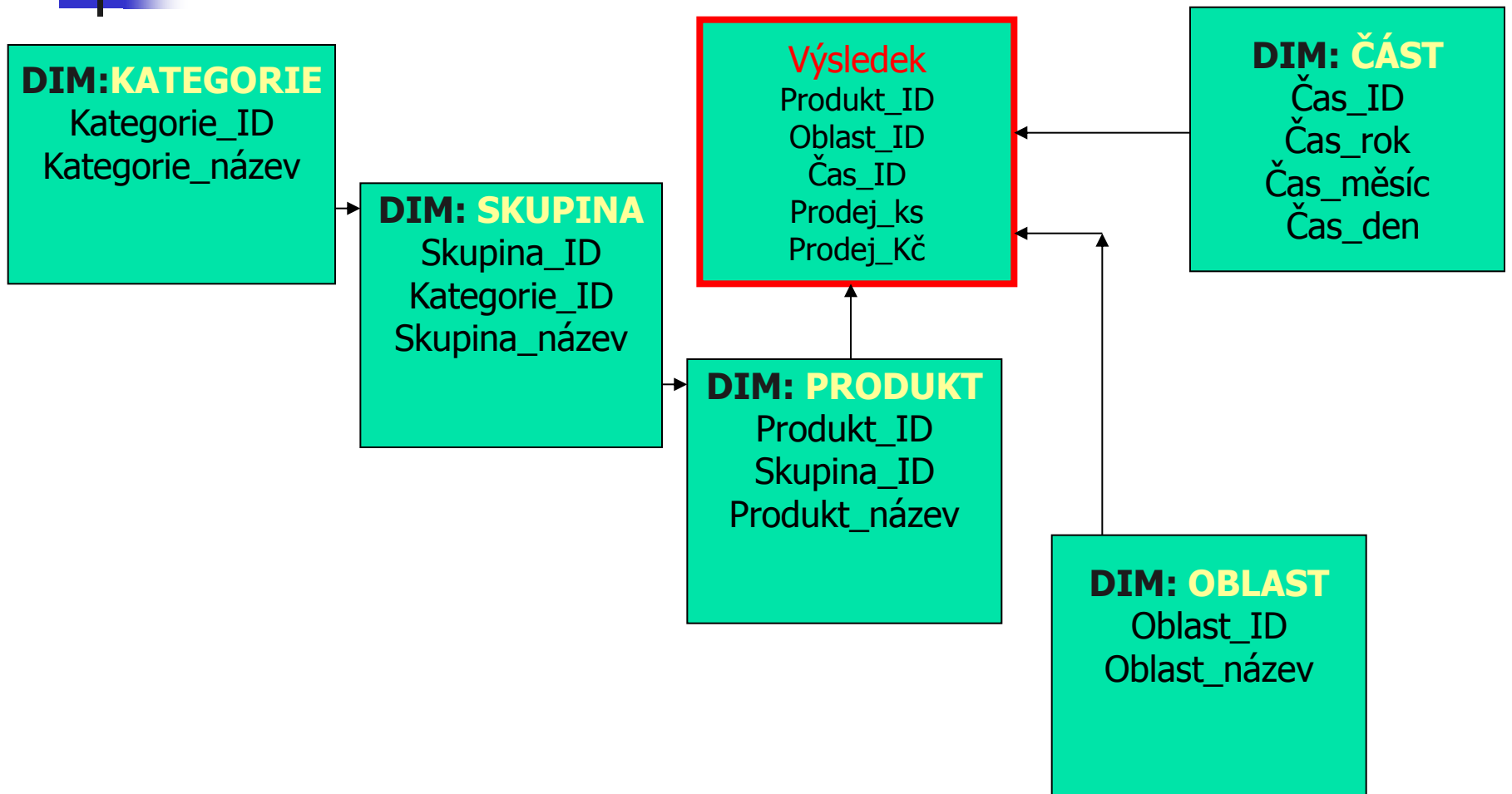
- [http://www.databaseanswers.org/designing\\_olap\\_cubes.htm](http://www.databaseanswers.org/designing_olap_cubes.htm)



# Relační dimenzionální model: STAR



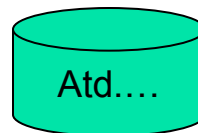
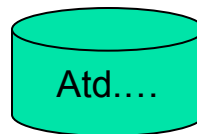
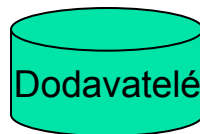
# Relační dimenzionální model: SNOWFLAKE



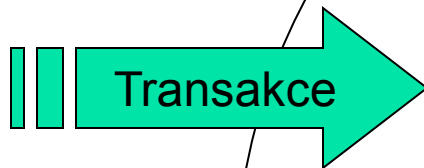
# Datový sklad



## Podniková DB



Kopie, ,  
organizace dat  
Sumarizace dat



## Datoví horníci :

- "Profíci" – vědí co chtějí
- "Výzkumníci" – nepředvídané výsledky





# Definice

---

- Datový sklad: základní komponenta BI
- Datové tržiště : subjektově orientované analytické DB- součást datového skladu
- Operativní datová úložiště : podpůrné analytické DB
- Dočasná úložiště dat : úložiště dat před jejich zpracování do databázových komponent řešení BI



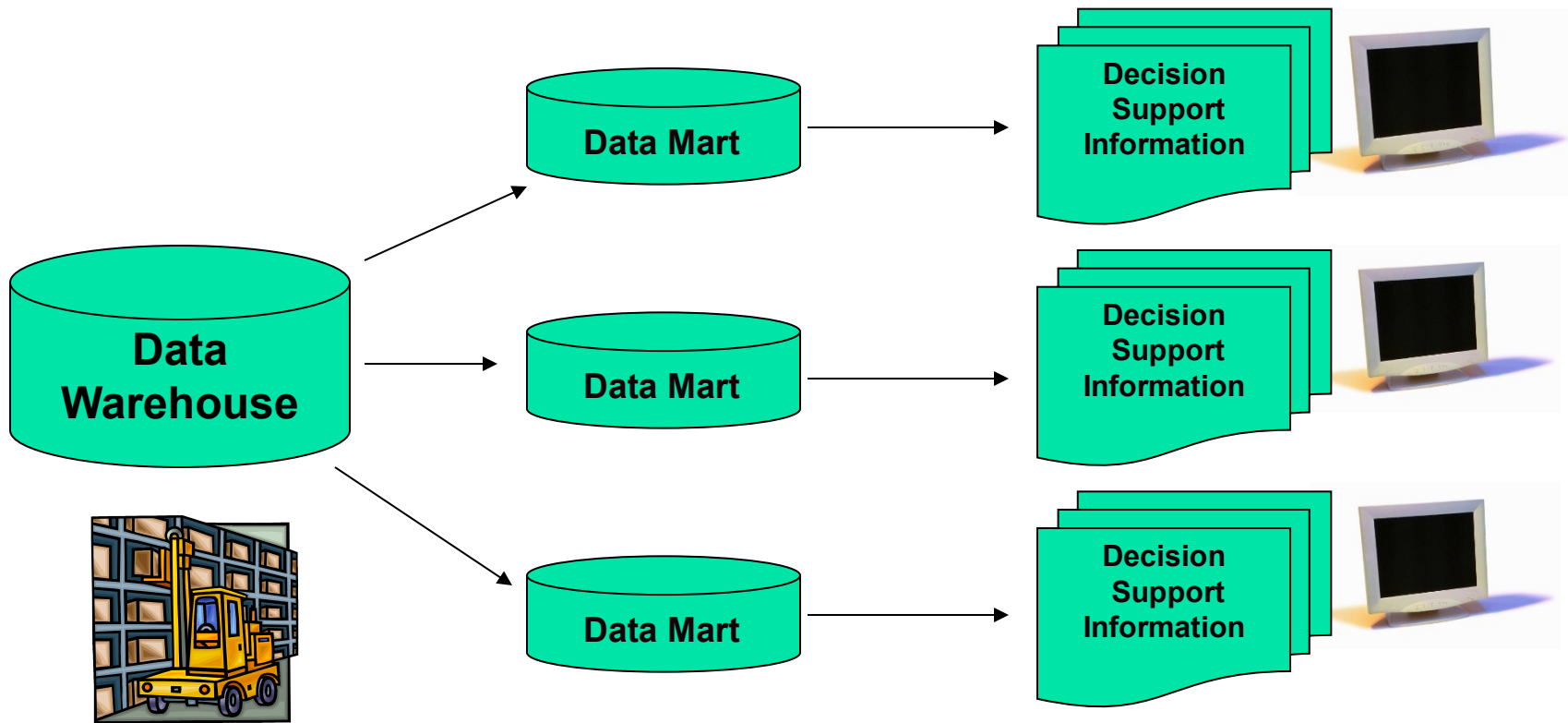


# Vrstvy pro analýzu dat

---

- Reporting : ad hoc dotazovací proces do DB komponent BI
- (ad hoc = k tomuto, za tímto účelem, pro tento jednotlivý případ)
- OLAP : pokročilé a dynamické analytické úlohy
  
- Data Mining (dolování dat) : sofistikovaná analýza většího množství dat
  
- Algoritmy pro dolování dat :
  - rozhodovací stromy
  - Neuronové sítě
  - Clustering a klasifikace

# Datový sklad->datové tržiště (anglická verze)





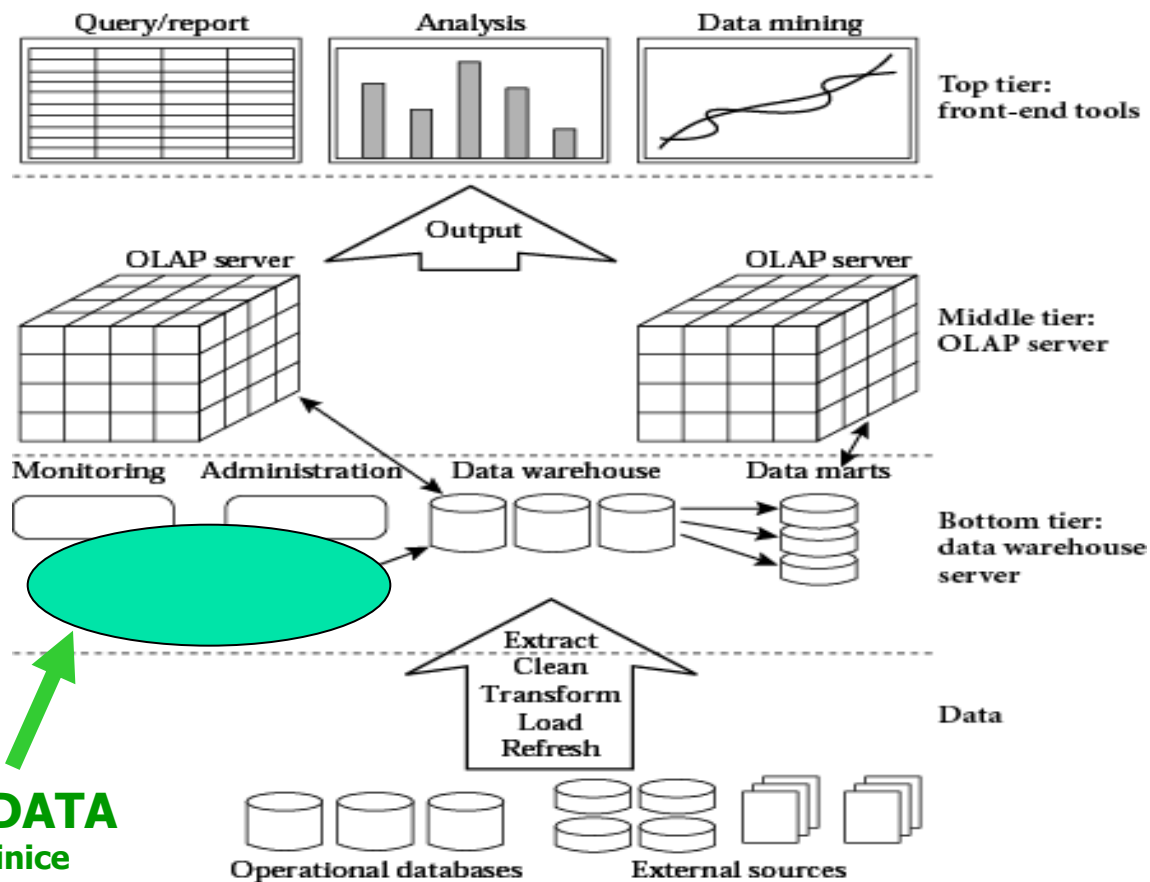
# Vysvětlení pojmu METADATA

---

**Metadata** jsou data o datech, kde pomocí předem definovaných dat s jasně danou a popsanou strukturou uchováváme informace o jiných datech.

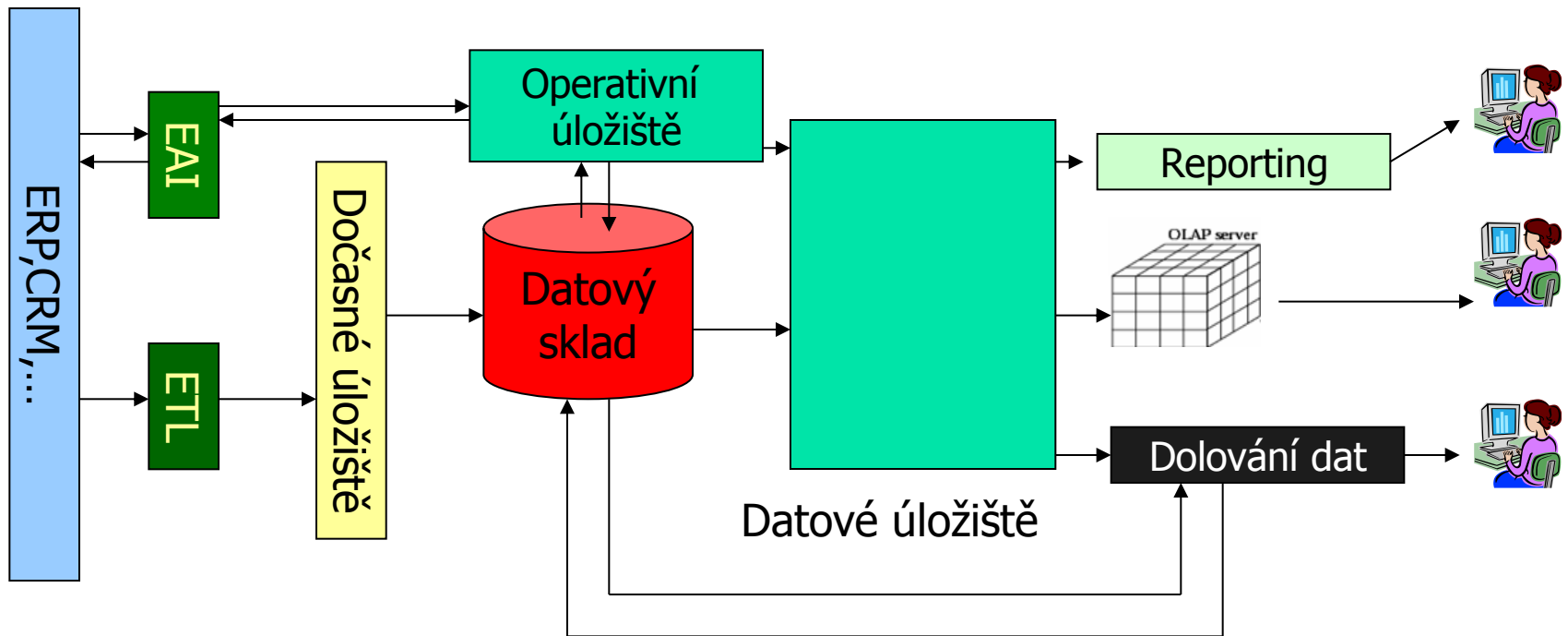
Typickým příkladem **metadat** jsou katalogizační záznamy v knihovnách, což byla jejich původní funkce.

# Architektura OLAP (anglická verze)



**METADATA**  
viz definice  
na předchozím snímku

# Hlavní komponenty BI a jejich vazby



Transformační nástroje (ETL : Extraction Transformation Loading)  
Integrační nástroje (EAI : Enterprise Application Integration)

# Dolování dat



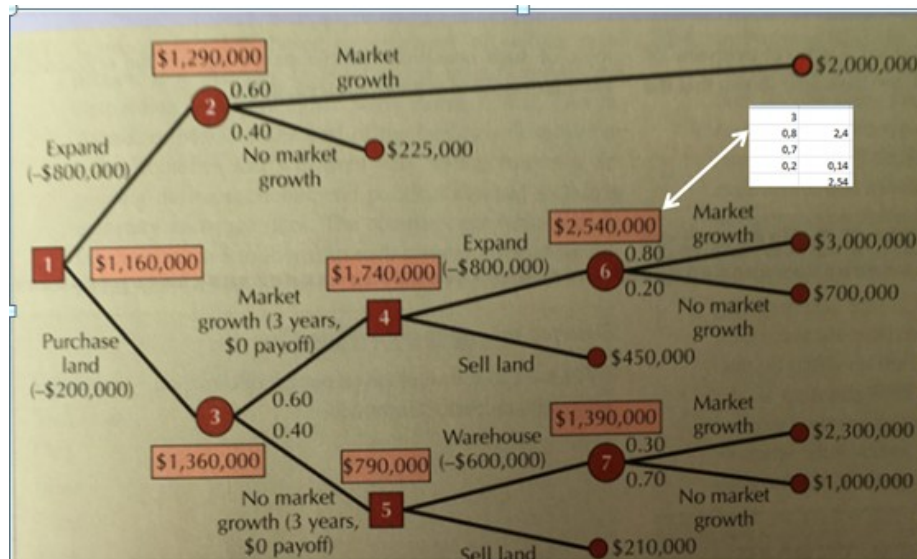
- Rozhodovací stromy
- Neuronové sítě
- Genetické algoritmy
- Clustering a klasifikace



# Dolování dat

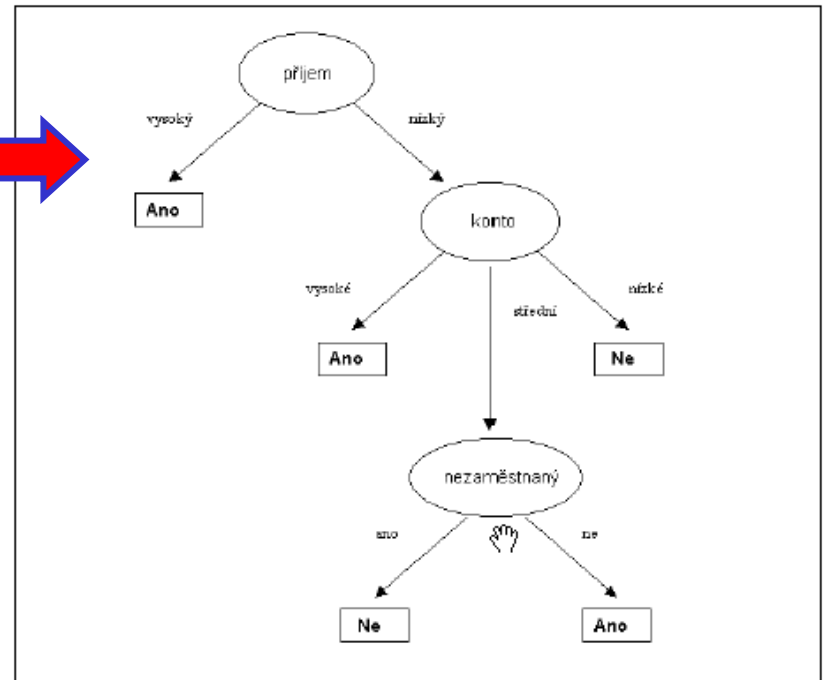


- Rozhodovací stromy (RS) - prediktivní model, který se zobrazuje v podobě stromu, kde každý uzel určuje kritérium pro následní rozvětvení. Strom rozděljuje veškerá zdrojová data do segmentů, kde každý list odpovídá určitému segmentu definovanému předešlými uzly. Data v jednom segmentu mají shodné vlastnosti.



# Příklad vytvoření RS

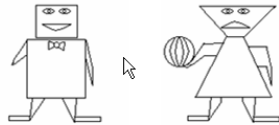
| klient | příjem | konto   | pohlaví | nezaměstnaný | úvěr |
|--------|--------|---------|---------|--------------|------|
| k1     | vysoký | vysoké  | žena    | ne           | ano  |
| k2     | vysoký | vysoké  | muž     | ne           | ano  |
| k3     | nízký  | nízké   | muž     | ne           | ne   |
| k4     | nízký  | vysoké  | žena    | ano          | ano  |
| k5     | nízký  | vysoké  | muž     | ano          | ano  |
| k6     | nízký  | nízké   | žena    | ano          | ne   |
| k7     | vysoký | nízké   | muž     | ne           | ano  |
| k8     | vysoký | nízké   | žena    | ano          | ano  |
| k9     | nízký  | střední | muž     | ano          | ne   |
| k10    | vysoký | střední | žena    | ne           | ano  |
| k11    | nízký  | střední | žena    | ano          | ne   |
| k12    | nízký  | nízký   | muž     | ne           | ano  |



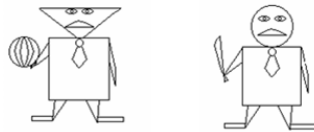
## Typy stromů :

- CART=Classification and Regression Trees  
(kriterium redukce směrodatné odchylky)
- CHAID =Chi-squared Automatic Interaction Detector

# Rozdělení postaviček podle atributů



přátelští

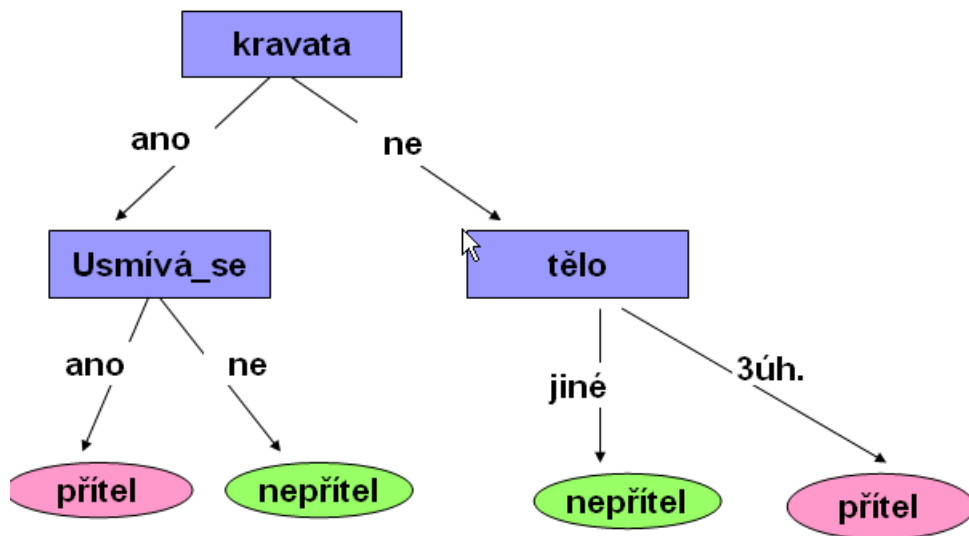


nepřátelští



| Hlava       | Úsměv | Ozdoba  | Tvar těla   | Předmět | Přátelský |
|-------------|-------|---------|-------------|---------|-----------|
| Kruh        | Ne    | Kravata | Čtverec     | Šavle   | NE        |
| Čtverec     | Ano   | Motýlek | Čtverec     | NIC     | ANO       |
| Kruh        | Ne    | Motýlek | Kruh        | Šavle   | ANO       |
| Trojúhelník | Ne    | Kravata | Čtverec     | Balon   | NE        |
| Kruh        | Ano   | NIC     | Trojúhelník | Květina | NE        |
| Trojúhelník | Ne    | NIC     | Trojúhelník | Balon   | ANO       |
| Trojúhelník | Ano   | Kravata | Kruh        | NIC     | NE        |
| Kruh        | Ano   | Kravata | Kruh        | NIC     | ANO       |

# Rozhodovací strom jako logický výraz



(Kravata=ano & usmívá\_se=ano) **V** (Kravata=ne & tělo=3úh.)



# OLAP databáze

---

- **OLAP** DB představují jednu nebo více souvisejících OLAP kostek
- **OLAP** kostka na rozdíl od datových skladů zahrnuje předzpracované agregace dat podle definovaných hierarchických struktur dimenzí a jejich kombinací
- Technologie **OLAP** má několik variant (uvádím zde pouze dvě z nich):
  - **MOLAP** - Multidimensional OLAP (speciální uložení v multidimenzionálních-binárních kostkách)
  - **ROLAP** – Relational OLAP (uloží data do relační DB)

# Datová pumpa

**Primární  
transakční systém  
(ERP, CRM,..)**



**Datová  
pumpa**



**Datový  
sklad**

# Datová pumpa (kritické místo celé aplikace)

Datová pumpa, nebo-li ETL nástroj umožňuje efektivní zpracování velkých objemů z různých zdrojů a jejich uložení do datového skladu. Každý ETL nástroj musí umět:

- zpracovávat různorodá data obvykle fyzicky umístěná na různých místech,
- navrhovat transformace pro přenos dat mezi různými datovými formáty



**Zpracování** = odstranění redundancí, agregace podle dimenzí, zapomínání dat  
**Zapomínání dat** = úmyslné odstranění nepotřebných dat z datového skladu

# Datová pumpa (kritické místo celé aplikace)

Datová pumpa = Extraction Transformation and Loading = ETL

