

Čistota dat

Základ datové kvality

Agenda

- Základní principy
- Data quality scorecard

Datová kvalita

Datová kvalita je proces detekce, reportingu a nápravy neplatných nebo chybějících hodnot v datech.

Před

Prevence:	10%
Detekce:	30%
Oprava:	60%

Po

Prevence:	45%
Detekce:	30%
Oprava:	25%

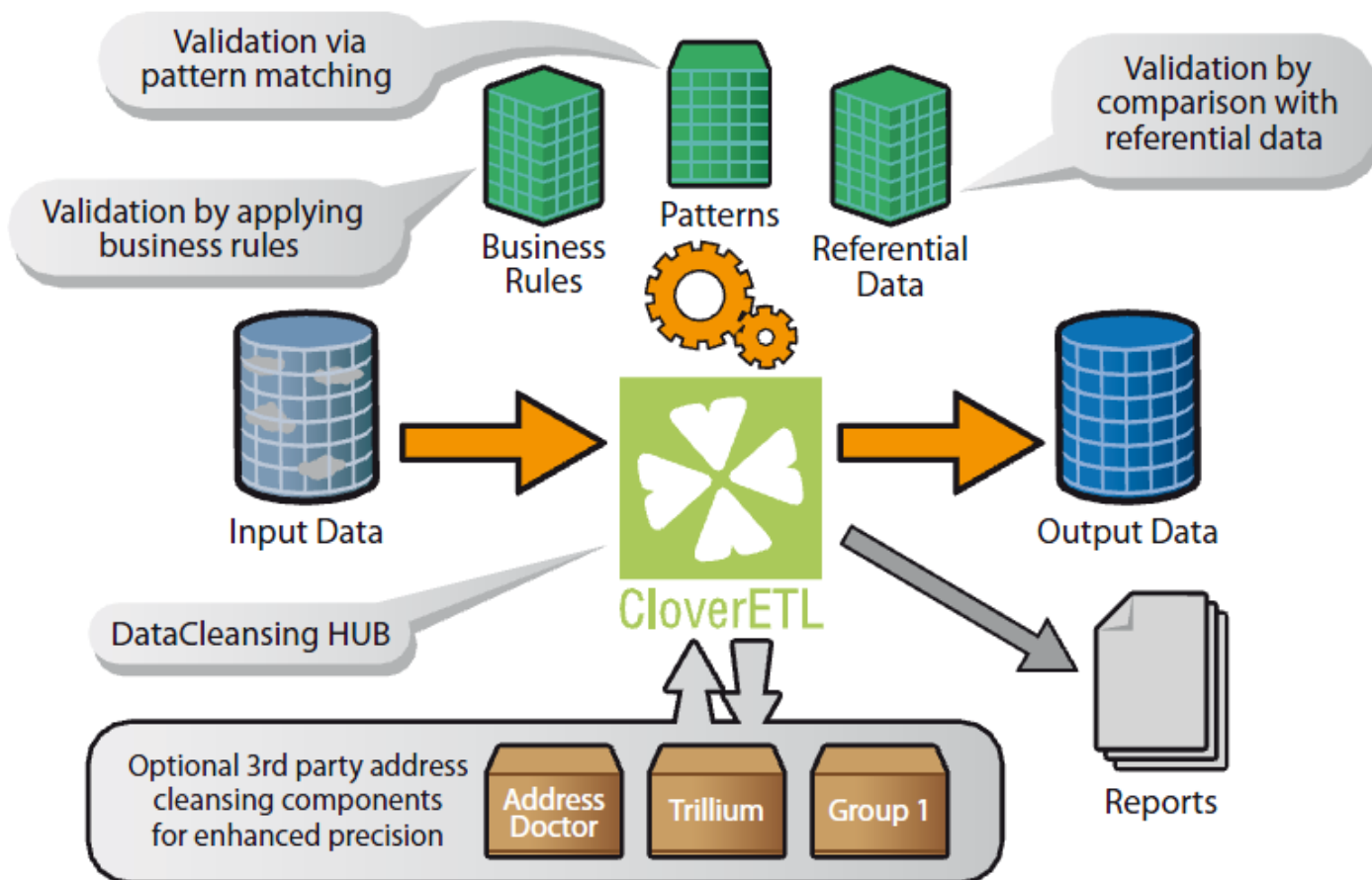
Základní principy

- Validace datového typu
 - Datum, číslo
- Validace formátu
 - Částka, měna, datum
- Validace kontrolní číslice
 - IČ, RČ, číslo účtu

Základní principy

- Porovnání s referenčními daty
 - Interní datové zdroje
 - CRM, LDAP, ERP
 - Veřejné datové zdroje
 - RES, ARES, registr dlužníků
 - Komerční datové zdroje
 - Trillium, AddressDoctor
 - Adresní body České pošty

Čištění dat

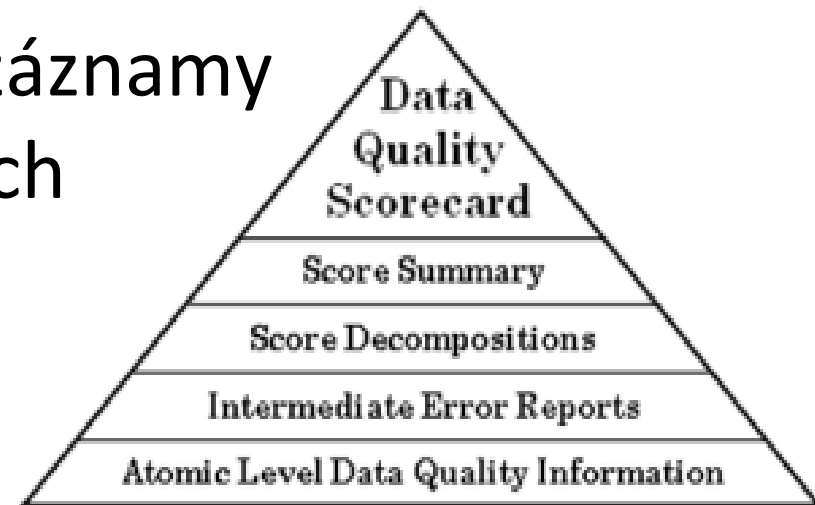


Proces datové kvality

- Implementace datové kvality je dvoufázový proces
 - Fáze 1 - Iniciální vyčištění
 - Fáze 2 - Průběžný monitoring, prevence a náprava
- Užitečným nástrojem je data quality scorecard

Data quality scorecard

- Užitečný analytický nástroj, který umožňuje sledovat finanční dopad špatných dat a odhadnout návratnost investic do DQ aktivit
- Důležitá pro rozhodování jsou zejména agregovaná skóre v horní části hierarchie
- Na nejnižší úrovni jsou záznamy o jednotlivých záznamech



Výběr skóre

- Obecná agregovaná skóre nejsou příliš vhodná
 - 6,3% všech záznamů v databázi obsahuje chybu
 - Je to dobrá, nebo špatná zpráva?
- Agregovat by se měla skóre, která sledují byznys cíle
 - 6,3% všech zákazníků obdrží při fakturaci slevu, na kterou nemá nárok
 - Finanční dopad se dá odhadnout!

Návratnost investic

- DQ scorecard je klíčový nástroj při posuzování ROI
- Pokud celkové náklady na korekci problémů nepřevýší celkové náklady (ztráty) plynoucí ze špatné kvality dat, bude mít akce pozitivní ekonomický přínos
- Z dlouhodobého hlediska bude přínos řádově vyšší

Návratnost investic

Ref. číslo	Zjištěný problém	Lokace	Business Proces	Dopad	Náklady
1	Chybné číslo kreditní karty	Node 5	Platební systém kreditních karet	Detekce	360.000
			Kontaktování zákazníka	Náprava Znovuzavedení	210.000 600.000
2	Chybná adresa	Node 1	Přímý marketing	Detekce Náprava	- 600.000
			Snížená návratnost	Náklady akvizice dat Ztracená příležitost	135.000 270.000
3	Chybný dodací list	Node 7	Expedice	Detekce Náprava	750.000 660.000
			Zákaznický servis	Plnění záruky Odchod zákazníka	1.290.000 1.500.000
					Σ 6.375.000

Klíčové otázky

- Analýza agregovaných skóre směrem dolů by měla odpovědět na následující klíčové otázky:
 - Jaký je dopad chyb ve vaší databázi na byznys procesy?
 - Jaké jsou zdroje a příčiny chyb ve vaší databázi?
 - Kde může být v databázi nalezeno nejvíce chyb?

Přínosy

- Zvýšení účinnosti marketingové kampaně
 - Vhodné adresy, vhodní lidé
- Systém včasného upozornění
 - Identifikace subjektů s problémy
- Zvýšení důvěryhodnosti a dobrého jména u zákazníků
 - Pošta a hovory správným lidem
 - Prevence opakovaného volání

Dotazy

Dotazy?
Děkuji za pozornost.