SUPER-X

SUPER-X – DATA WAREHOUSING

Group 6: Procurement

| 14th February 2018

SUPER-X

CONTENT

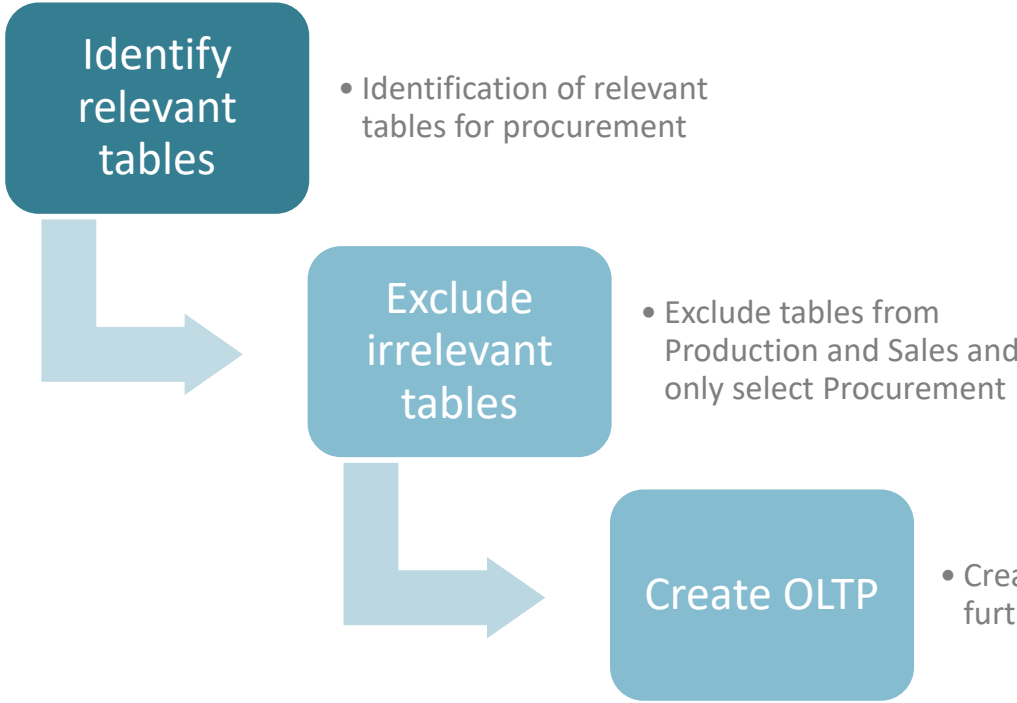
- REQUIREMENT ANALYSIS
- ANALYSIS OF DATA SOURCES
- CONCEPTUAL DESIGN
- PROOF-OF-CONCEPT
- PROCESS INTELLIGENCE
- BUSINESS RECOMMENDATIONS

SUPER-X

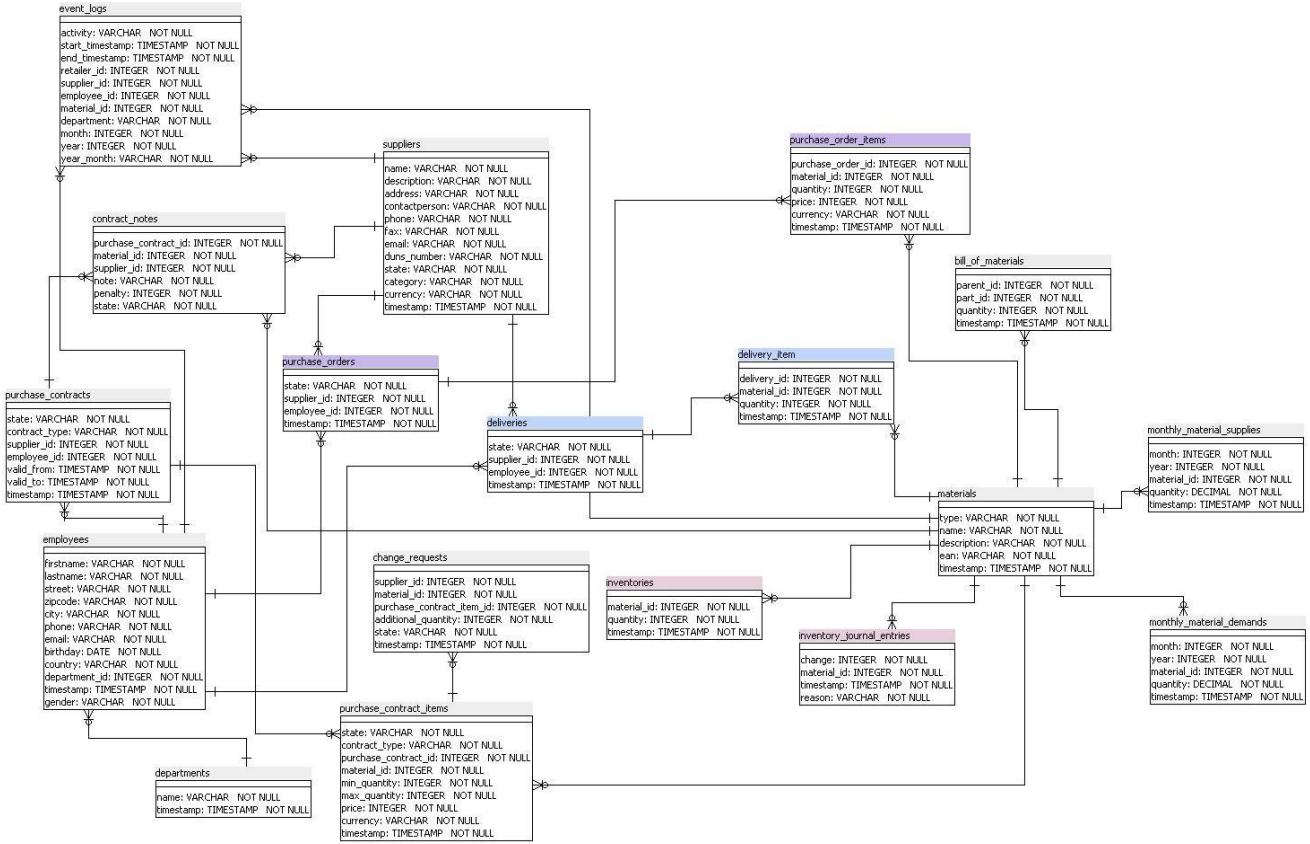
CONTENT

- **REQUIREMENT ANALYSIS**
- ANALYSIS OF DATA SOURCES
- CONCEPTUAL DESIGN
- PROOF-OF-CONCEPT
- PROCESS INTELLIGENCE
- BUSINESS RECOMMENDATIONS

REQUIREMENT ANALYSIS



OLTP SCHEMA


event_logs
activity: VARCHAR NOT NULL
start_timestamp: TIMESTAMP NOT NULL
end_timestamp: TIMESTAMP NOT NULL
retailer_id: INTEGER NOT NULL
supplier_id: INTEGER NOT NULL
employee_id: INTEGER NOT NULL
material_id: INTEGER NOT NULL
department: VARCHAR NOT NULL
month: INTEGER NOT NULL
year: INTEGER NOT NULL
year_month: VARCHAR NOT NULL
suppliers
name: VARCHAR NOT NULL
description: VARCHAR NOT NULL
address: VARCHAR NOT NULL
contactperson: VARCHAR NOT NULL
phone: VARCHAR NOT NULL
fax: VARCHAR NOT NULL
email: VARCHAR NOT NULL
duns_number: VARCHAR NOT NULL
state: VARCHAR NOT NULL
category: VARCHAR NOT NULL
currency: VARCHAR NOT NULL
timestamp: TIMESTAMP NOT NULL
purchase_order_items
purchase_order_id: INTEGER NOT NULL
material_id: INTEGER NOT NULL
quantity: INTEGER NOT NULL
price: INTEGER NOT NULL
currency: VARCHAR NOT NULL
timestamp: TIMESTAMP NOT NULL
contract_notes
purchase_contract_id: INTEGER NOT NULL
material_id: INTEGER NOT NULL
supplier_id: INTEGER NOT NULL
note: VARCHAR NOT NULL
penalty: INTEGER NOT NULL
state: VARCHAR NOT NULL
bill_of_materials
parent_id: INTEGER NOT NULL
part_id: INTEGER NOT NULL
quantity: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
purchase_orders
state: VARCHAR NOT NULL
supplier_id: INTEGER NOT NULL
employee_id: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
delivery_item
delivery_id: INTEGER NOT NULL
material_id: INTEGER NOT NULL
quantity: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
purchase_contracts
state: VARCHAR NOT NULL
contract_type: VARCHAR NOT NULL
supplier_id: INTEGER NOT NULL
employee_id: INTEGER NOT NULL
valid_from: TIMESTAMP NOT NULL
valid_to: TIMESTAMP NOT NULL
timestamp: TIMESTAMP NOT NULL
deliveries
state: VARCHAR NOT NULL
supplier_id: INTEGER NOT NULL
employee_id: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
monthly_material_supplies
month: INTEGER NOT NULL
year: INTEGER NOT NULL
material_id: INTEGER NOT NULL
quantity: DECIMAL NOT NULL
timestamp: TIMESTAMP NOT NULL
materials
type: VARCHAR NOT NULL
name: VARCHAR NOT NULL
description: VARCHAR NOT NULL
ean: VARCHAR NOT NULL
timestamp: TIMESTAMP NOT NULL
employees
firstname: VARCHAR NOT NULL
lastname: VARCHAR NOT NULL
street: VARCHAR NOT NULL
zipcode: VARCHAR NOT NULL
city: VARCHAR NOT NULL
phone: VARCHAR NOT NULL
email: VARCHAR NOT NULL
birthday: DATE NOT NULL
country: VARCHAR NOT NULL
department_id: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
gender: VARCHAR NOT NULL
change_requests
supplier_id: INTEGER NOT NULL
material_id: INTEGER NOT NULL
purchase_contract_item_id: INTEGER NOT NULL
additional_quantity: INTEGER NOT NULL
state: VARCHAR NOT NULL
timestamp: TIMESTAMP NOT NULL
inventories
material_id: INTEGER NOT NULL
quantity: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
inventory_journal_entries
change: INTEGER NOT NULL
material_id: INTEGER NOT NULL
timestamp: TIMESTAMP NOT NULL
reason: VARCHAR NOT NULL
monthly_material_demands
month: INTEGER NOT NULL
year: INTEGER NOT NULL
material_id: INTEGER NOT NULL
quantity: DECIMAL NOT NULL
timestamp: TIMESTAMP NOT NULL
purchase_contract_items
state: VARCHAR NOT NULL
contract_type: VARCHAR NOT NULL
purchase_contract_id: INTEGER NOT NULL
material_id: INTEGER NOT NULL
min_quantity: INTEGER NOT NULL
max_quantity: INTEGER NOT NULL
price: INTEGER NOT NULL
currency: VARCHAR NOT NULL
timestamp: TIMESTAMP NOT NULL
departments
name: VARCHAR NOT NULL
timestamp: TIMESTAMP NOT NULL

KPIs

Questions	Business Requirement	Importance	HIgh level Entities	Measures
Q1	What is the average quantity bought for each material and material type per month and year?	Medium	Material, Material Type, Month, Year	Quantity bought
Q2	What is the total quantity bought for each material and material type per month and year?	High	Material, Material Type, Month, Year	Quantity bought
Q3	What is the highest and lowest price per material per month and year?	Medium	Material, Month, Year	Price
Q4	What is the total order value per month and year?	High	Month, Year	Price * Quantity
Q5	What is the total order value per supplier and supplier category per month and year	High	Supplier, Supplier Category, Month, Year	Price * Quantity
Q6	What is the total quantity ordered per country?	Low	Country	Quantity
Q7	What is the order volume per supplier and supplier category per month and year?	High	Supplier, Supplier Category, Month, Year	Quantity
Q8	What is the number of suppliers per material and supplier category per month and year?	HIgh	Supplier, Supplier Category, Material, Month, Year	
Q9	What is the order volume per supplier and country per month and year?	Medium	Supplier, Supplier Category, Country, Month, Year	Quantity
Q10	What is the order value per supplier and country per month and year?	Medium	Supplier, Supplier Category, Country, Month, Year	Price * Quantity

SUPER-X

CONTENT

- REQUIREMENT ANALYSIS
- **ANALYSIS OF DATA SOURCES**
- CONCEPTUAL DESIGN
- PROOF-OF-CONCEPT
- PROCESS INTELLIGENCE
- BUSINESS RECOMMENDATIONS

DATA PROFILING

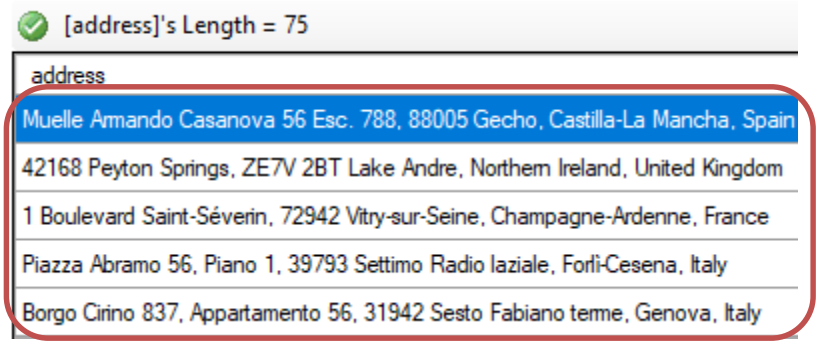
Types of data profiles used:

- Candidate Key Profiles
- Column Length Distribution Profiles
- Column Null Ratio Profiles
- Column Pattern Profiles
- Column Statistics Profiles
- Column Value Distribution Profiles
- Functional Dependency Profiles

DATA CLEANSING: PROBLEM

Example: Supplier table

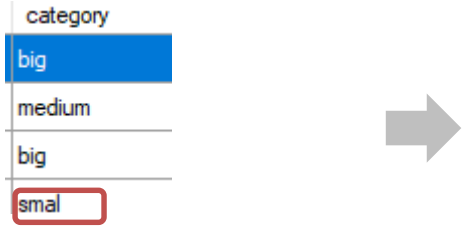
- Combined field for the whole address:



Microsoft Excel

street	zipcode	city	region	country
Emmapark 597 I	5326	NH Maas aan de Ijssel	Limburg	Netherlands
Am Alten Schafstall 85b	70119	Ost Fabianland	Hamburg	Deutschland
Rotonda Costa 2	41981	Sesto Elga	Varese	Italy
Meckhofer Feld 18	14490	Jarosburg	Bremen	Deutschland
498 Rutherford Row	M1O-5I8	Tremblaymouth	Manitoba	Canada
191 Leannon Ville	U3R3R6	East Savanna	Ontario	Canada
Glorieta Gabriela Cardenas 85	99566	Almería	Región de Murcia	Spain
Vriesplantsoen 888	6760-QK	Oud Annesluus	Limburg	Netherlands
990 Patsy View	89563-9582	Lake Samson	North Dakota	USA
97 Quai de la Harpe	94423	Neuilly-sur-Seine	Basse-Normandie	France

- Category “smal”:



DATA CLEANSING: SOLUTION

```
INSERT INTO [DataMart_NewSuperX].[dbo].[DimPurchaseOrder]
SELECT DISTINCT purchase_order_id, [state] = purchase_orders.state,
[supplier_name] = suppliers.name,
[supplier_category] = CASE WHEN category='smal' THEN 'small' ELSE category END,
[supplier_country] = CASE WHEN right(address, CHARINDEX(' ', REVERSE(address))-1)='U.S.A.' THEN 'USA'
WHEN right(address, CHARINDEX(' ', REVERSE(address))-1)='Deutschland' THEN 'Germany'
WHEN right(address, CHARINDEX(' ', REVERSE(address))-1)='Kingdom' THEN 'United Kingdom'
ELSE right(address, CHARINDEX(' ', REVERSE(address))-1) END,
[employee_name] = concat(employees.firstname, ' ', employees.lastname),
[effective_date] = cast(NewSuperX.dbo.purchase_orders.timestamp as date),
[current_flag] = 1
  FROM NewSuperX.dbo.purchase_order_items
  join NewSuperX.dbo.purchase_orders ON purchase_order_items.purchase_order_id = purchase_orders.id
  join NewSuperX.dbo.employees ON purchase_orders.employee_id = employees.id
  join NewSuperX.dbo.suppliers ON purchase_orders.supplier_id = suppliers.id;
go
```

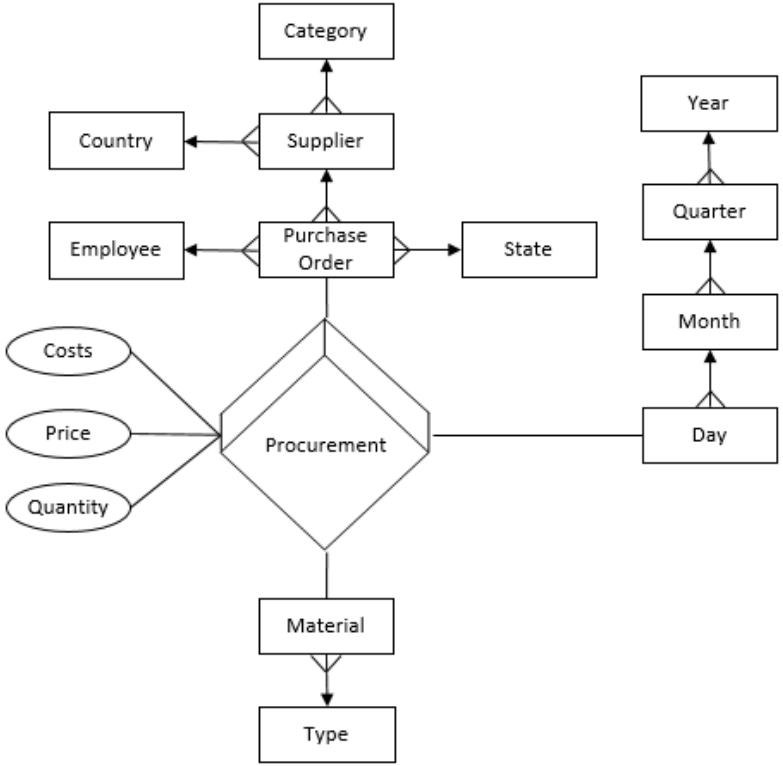
SUPER-X

CONTENT

- REQUIREMENT ANALYSIS
- ANALYSIS OF DATA SOURCES
- **CONCEPTUAL DESIGN**
- PROOF-OF-CONCEPT
- PROCESS INTELLIGENCE
- BUSINESS RECOMMENDATIONS

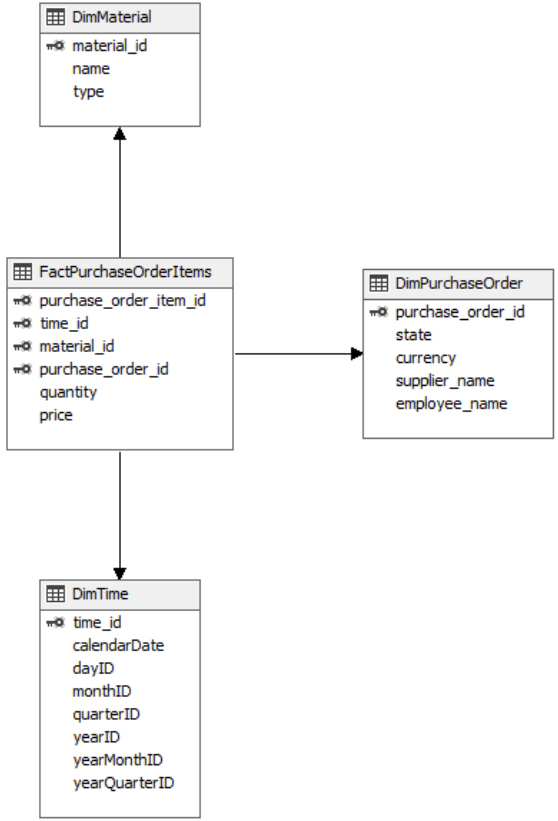
CONCEPTUAL DESIGN

MER diagram:



LOGICAL DESIGN

STAR schema:

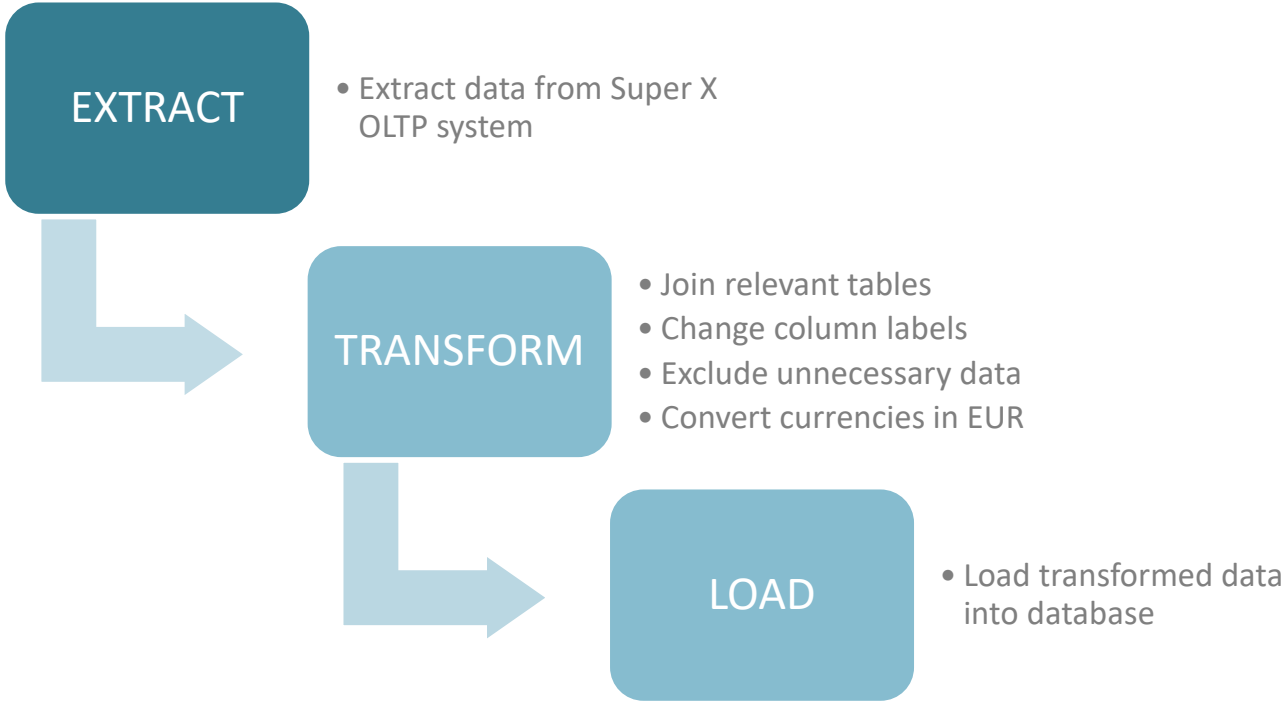


SUPER-X

CONTENT

- REQUIREMENT ANALYSIS
- ANALYSIS OF DATA SOURCES
- CONCEPTUAL DESIGN
- **PROOF-OF-CONCEPT**
- PROCESS INTELLIGENCE
- BUSINESS RECOMMENDATIONS

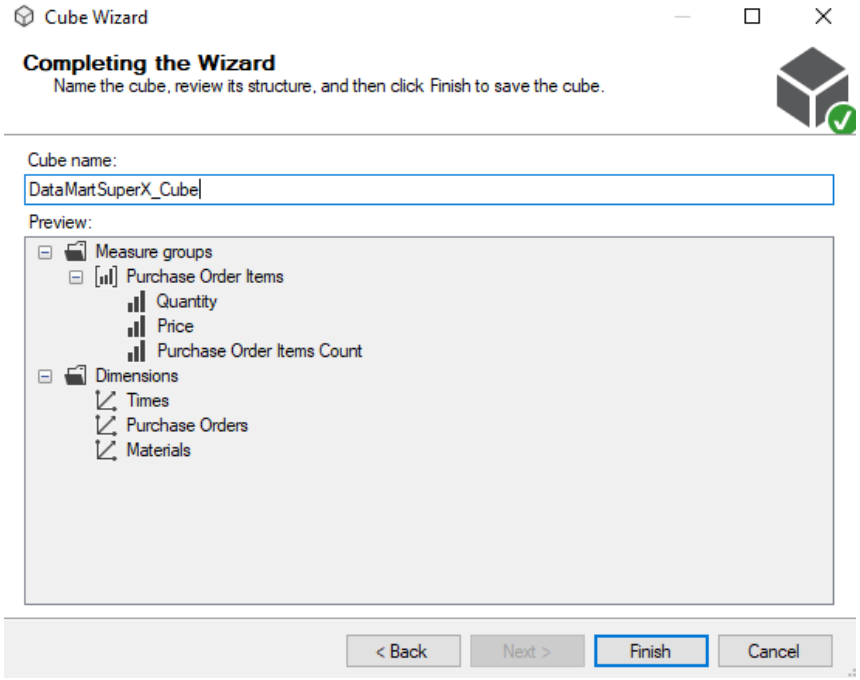
ETL PROCESS

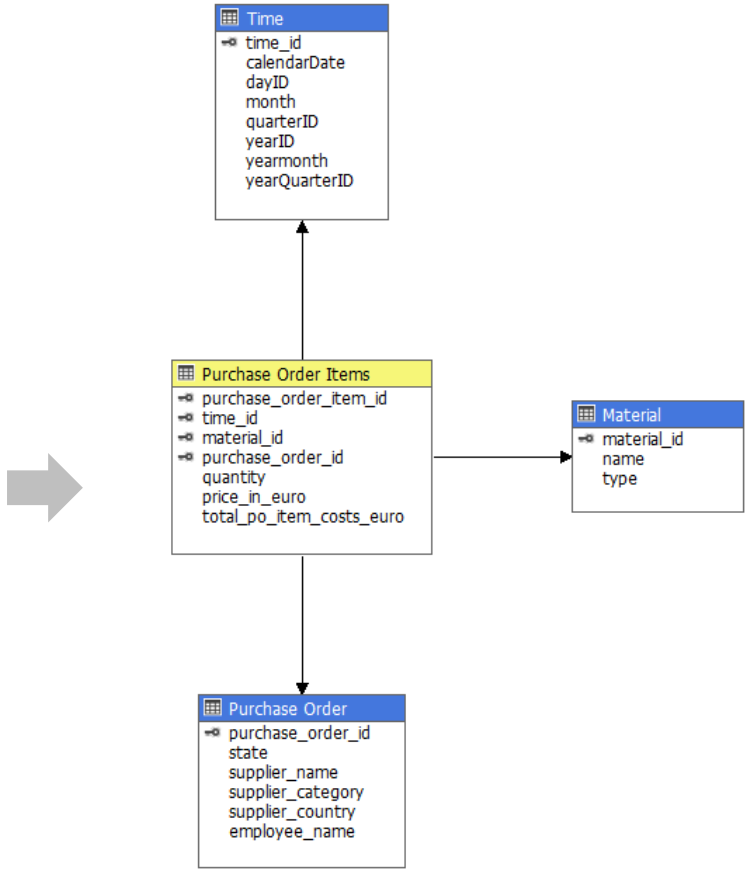

EXTRACT
• Extract data from Super X OLTP system
TRANSFORM
• Join relevant tables
• Change column labels
• Exclude unnecessary data
• Convert currencies in EUR
LOAD
• Load transformed data into database

ETL PROCESS

```
/* data imported into SQL SERVER*/
USE DataMart_NewSuperX;
GO

-- Dim Time
CREATE TABLE DimTime (
time_id int NOT NULL CONSTRAINT [pkDimTime] PRIMARY KEY,
calendarDate Date NOT NULL,
dayID int NOT NULL,
month nvarchar(50) NOT NULL,
quarterID int NOT NULL,
yearID int NOT NULL,
yearmonth nvarchar(7) NOT NULL,
yearQuarterID nvarchar(7) NOT NULL,
effective_date Date NOT NULL,
current_flag bit NOT NULL
);
go

-- Dim Material
CREATE TABLE DimMaterial (
material_id int NOT NULL CONSTRAINT [pkDimMaterial] PRIMARY KEY,
name nvarchar(50) NOT NULL,
type nvarchar(50) NOT NULL,
effective_date Date NOT NULL,
current_flag bit NOT NULL
);
go

-- Dim Purchase Order
CREATE TABLE DimPurchaseOrder (
purchase_order_id int NOT NULL CONSTRAINT [pkDimPurchaseOrder] PRIMARY KEY,
state nvarchar(50) NOT NULL,
supplier_name nvarchar(50) NOT NULL,
supplier_category nvarchar(50) NOT NULL,
supplier_country nvarchar(50) NOT NULL,
employee_name nvarchar(50) NOT NULL,
effective_date Date NOT NULL,
current_flag bit NOT NULL
);
go

-- Fact table Purchase Order Items
CREATE TABLE FactPurchaseOrderItems (
    purchase_order_item_id int NOT NULL, --PK
    time_id int, -- PK, FK1
    material_id int, -- PK, FK2
    purchase_order_id int, -- PK, FK3
    quantity int,
    price_in_euro money,
    total_po_item_costs_euro money
    CONSTRAINT [pkFactPurchaseOrderItems] PRIMARY KEY (purchase_order_item_id, time_id, material_id, purchase_order_id)
);
go

-- FK constraints
ALTER TABLE dbo.FactPurchaseOrderItems
ADD CONSTRAINT fkFactToDimTime FOREIGN KEY (time_id) REFERENCES dbo.DimTime (time_id);
go

ALTER TABLE dbo.FactPurchaseOrderItems
ADD CONSTRAINT fkFactToDimMaterial FOREIGN KEY (material_id) REFERENCES dbo.DimMaterial (material_id);
go

ALTER TABLE dbo.FactPurchaseOrderItems
ADD CONSTRAINT fkFactToDimPurchaseOrder FOREIGN KEY (purchase_order_id) REFERENCES dbo.DimPurchaseOrder (purchase_order_id);
go

-- INSERT DATA
INSERT INTO [DataMart_NewSuperX].[dbo].[DimTime]
SELECT distinct [time_id] = concat(YEAR(timestamp), MONTH(timestamp), DAY(timestamp)),
 [calendarDate] = cast(timestamp as date),
 [dayID] = DAY(timestamp),
 [month] = DATENAME(month,timestamp),
[quarterID] = DATEPART(quarter, timestamp),
[yearID] = YEAR(timestamp),
[yearmonth] = concat(YEAR(timestamp),'-', MONTH(timestamp)),
[yearQuarterID] = concat(YEAR(timestamp),'-', DATEPART(quarter, timestamp)),
[effective_date] = cast(timestamp as date),
[current_flag] = 1
from NewSuperX.dbo.purchase_order_items;
go

INSERT INTO [DataMart_NewSuperX].[dbo].[DimMaterial]
SELECT DISTINCT material_id, name, type,
[effective_date] = cast(NewSuperX.dbo.materials.timestamp as date),
[current_flag] = 1
  FROM NewSuperX.dbo.purchase_order_items
  join NewSuperX.dbo.materials ON purchase_order_items.material_id = materials.id;
go

INSERT INTO [DataMart_NewSuperX].[dbo].[DimPurchaseOrder]
SELECT DISTINCT purchase_order_id, [state] = purchase_orders.state,
[supplier_name] = suppliers.name,
[supplier_category] = CASE WHEN category='small' THEN 'small' ELSE category END,
[supplier_country] = CASE WHEN right(address, CHARINDEX(' ', REVERSE(address))-1)='U.S.A.' THEN 'USA'
WHEN right(address, CHARINDEX(' ', REVERSE(address))-1)='Deutschland' THEN 'Germany'
WHEN right(address, CHARINDEX(' ', REVERSE(address))-1)='Kingdom' THEN 'United Kingdom'
ELSE right(address, CHARINDEX(' ', REVERSE(address))-1) END,
[employee_name] = concat(employees.firstname, ' ', employees.lastname),
[effective_date] = cast(NewSuperX.dbo.purchase_orders.timestamp as date),
[current_flag] = 1
  FROM NewSuperX.dbo.purchase_order_items
  join NewSuperX.dbo.purchase_orders ON purchase_order_items.purchase_order_id = purchase_orders.id
  join NewSuperX.dbo.employees ON purchase_orders.employee_id = employees.id
  join NewSuperX.dbo.suppliers ON purchase_orders.supplier_id = suppliers.id;
go

CREATE OR ALTER VIEW cleaned_po_items
as with notnullcurrencies (purchase_order_id, nncurrency)
   as (SELECT distinct purchase_order_id, currency as nncurrency from NewSuperX.dbo.purchase_order_items where currency is not null)
    SELECT [purchase_order_item_id] = id,
  [time_id] = concat(YEAR(timestamp), MONTH(timestamp), DAY(timestamp)),
  material_id, [purchase_order_id] = purchase_order_items.purchase_order_id, quantity,
   case when purchase_order_items.currency is null then nncurrency else purchase_order_items.currency end as currency
     FROM NewSuperX.dbo.purchase_order_items
     join notnullcurrencies on notnullcurrencies.purchase_order_id = purchase_order_items.purchase_order_id;
go

INSERT INTO [DataMart_NewSuperX].[dbo].[FactPurchaseOrderItems]
SELECT c.purchase_order_item_id, c.time_id, c.material_id, c.purchase_order_id, c.quantity,
[price_in_euro] =
  CASE WHEN c.currency='CAD' THEN price*0.64
  WHEN c.currency='USD' THEN price*0.81
  WHEN c.currency='PLN' THEN price*0.24
  WHEN c.currency='GBP' THEN price*1.12
  ELSE price END,
  [total_po_item_costs_euro] = c.quantity*CASE WHEN c.currency='CAD' THEN price*0.64
  WHEN c.currency='USD' THEN price*0.81
  WHEN c.currency='PLN' THEN price*0.24
  WHEN c.currency='GBP' THEN price*1.12
  ELSE price END
FROM cleaned_po_items c
join [NewSuperX].[dbo].[purchase_order_items] on purchase_order_items.id = c.purchase_order_item_id;
```

IMPLEMENTED CUBE



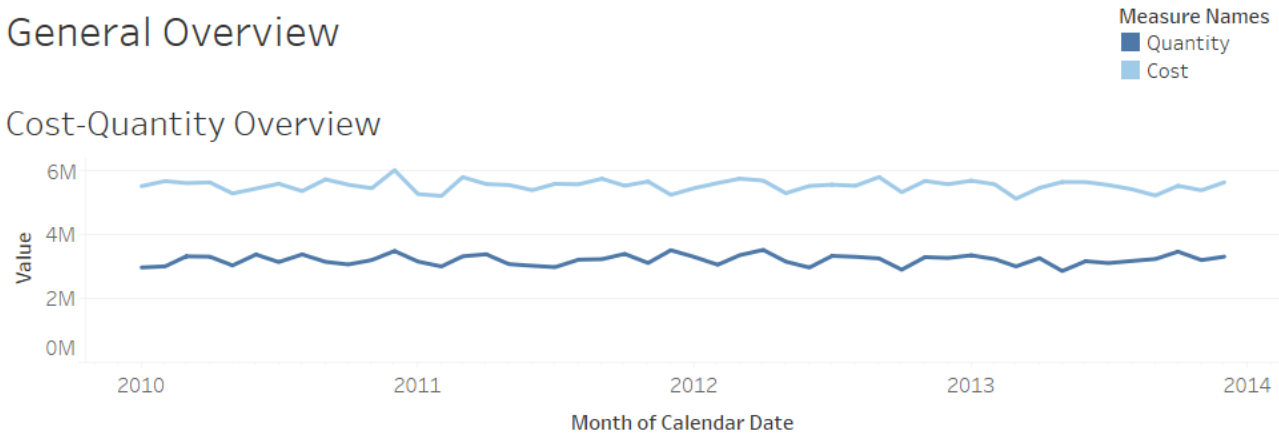


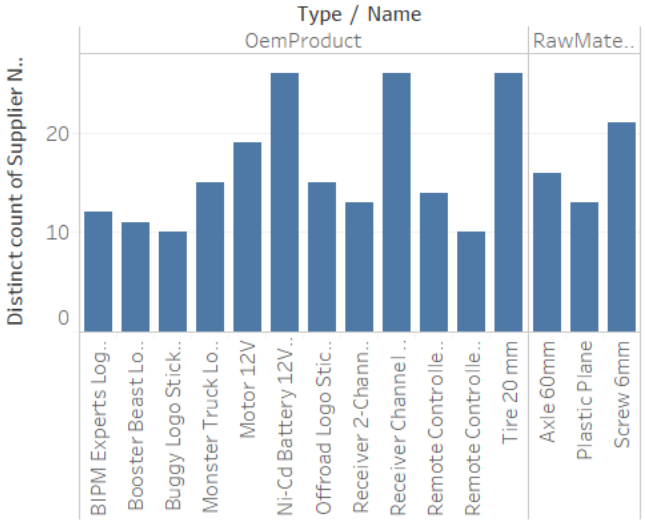
VISUALIZATION

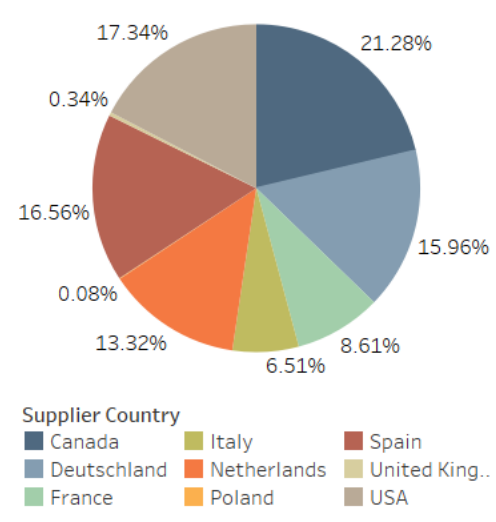
Advantages:

+ Userfriendly Interface
+ Vivid visualizations
+ In-Memory Architecture
+ Quick Insights

General Overview







Suppliers Dashboard

Choose Number of Suppliers:
27

Filters

Year Filter
2013

Month Filter
All

Supplier Country Filter
All

Supplier Category Filter
all

of Suppliers in Category

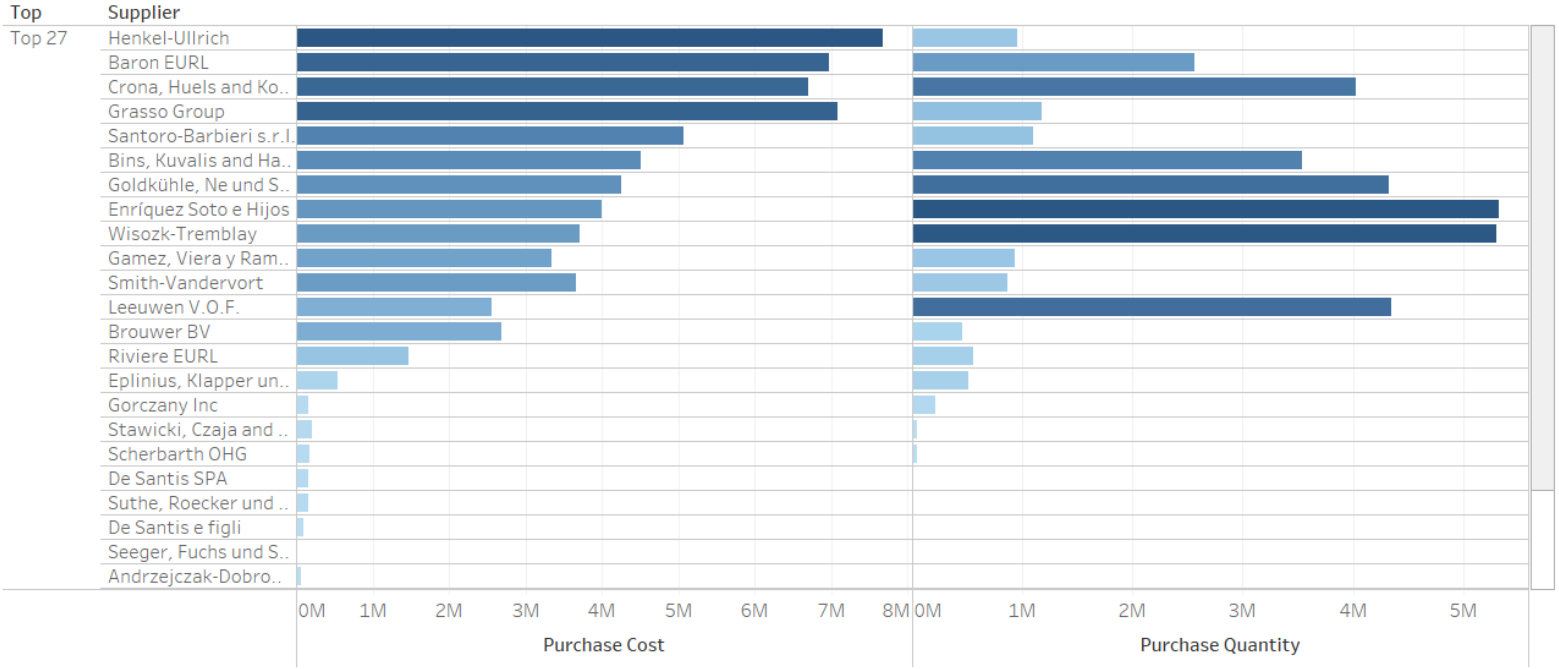
Supplier Category

Legend

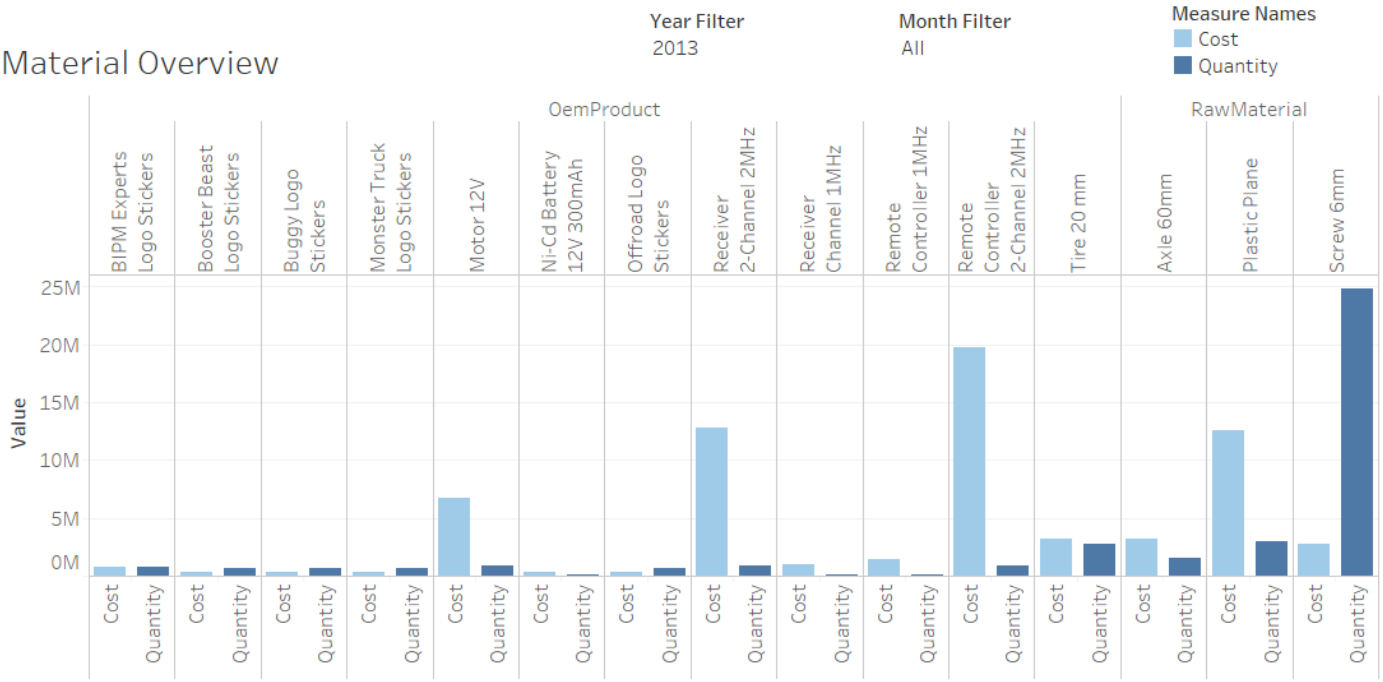
Purchase Cost
22,356 7,674,707

Purchase Quantity
6,298 5,325,612

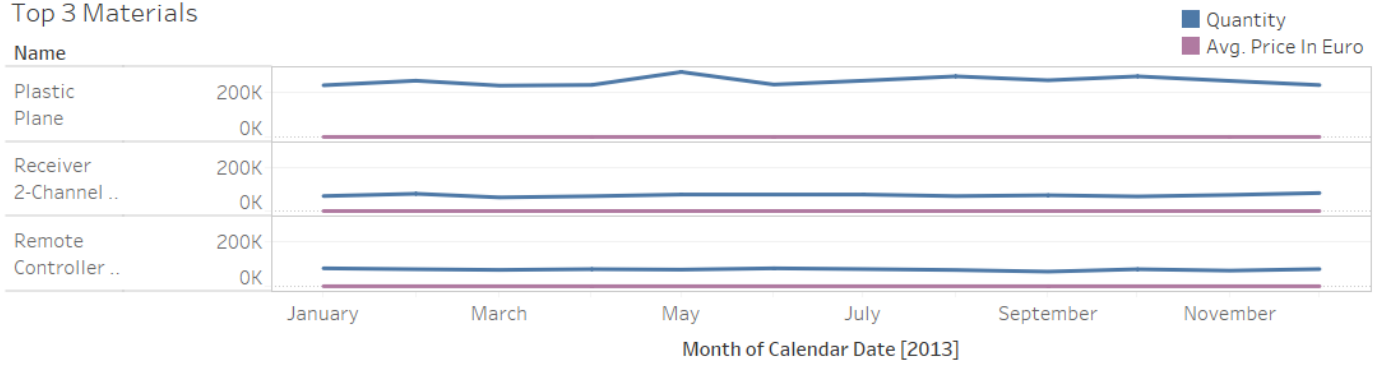
Top Suppliers Overview



Material Overview



Top 3 Materials

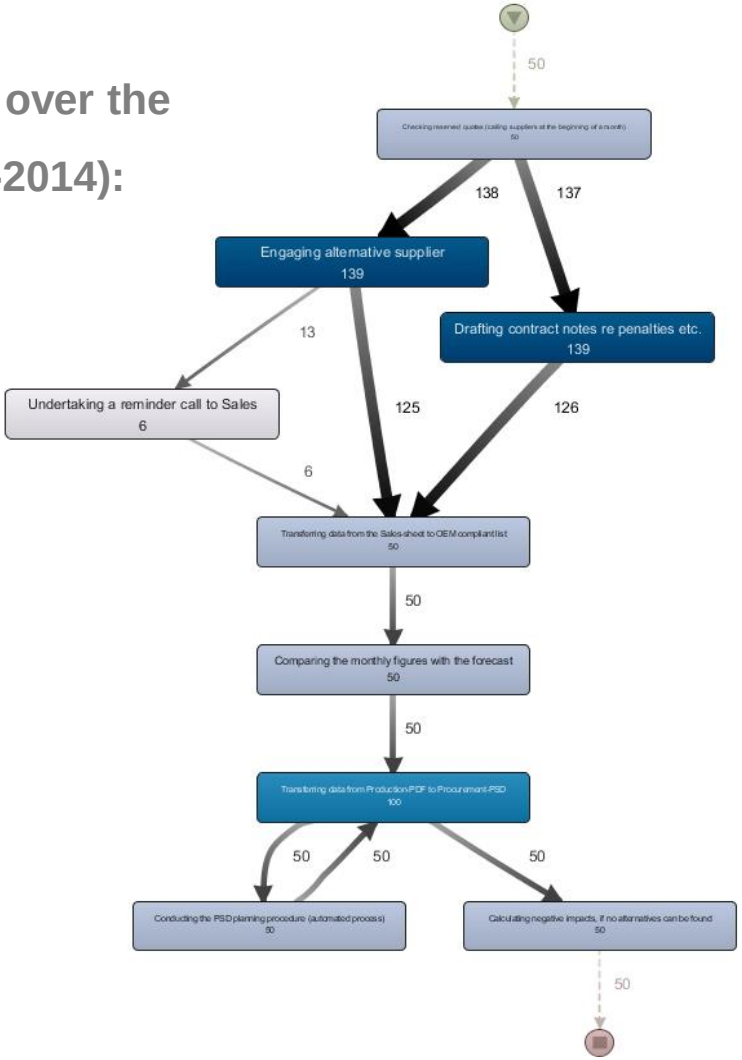


SUPER-X

CONTENT

- REQUIREMENT ANALYSIS
- ANALYSIS OF DATA SOURCES
- CONCEPTUAL DESIGN
- PROOF-OF-CONCEPT
- **PROCESS INTELLIGENCE**
- BUSINESS RECOMMENDATIONS

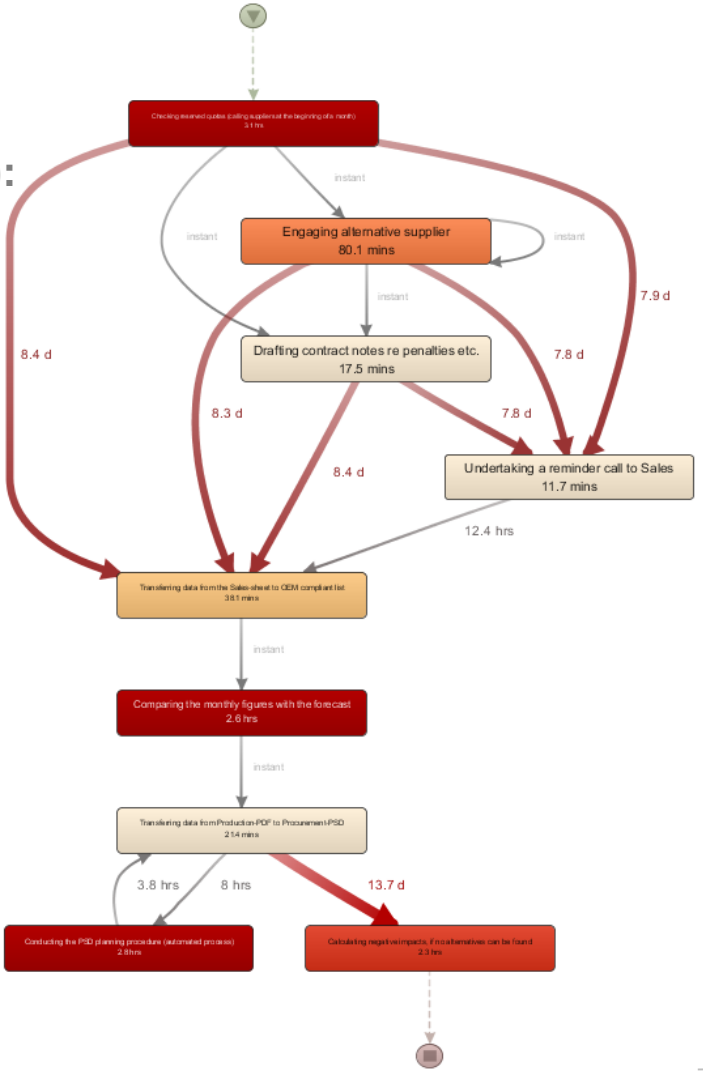
Frequency of tasks over the whole period (2010-2014):



Findings:

- Inefficiency
- Repetition

Mean duration of tasks over the whole period (2010-2014):



Findings:

- Waiting times
- Time consuming tasks

SUPER-X

CONTENT

- REQUIREMENT ANALYSIS
- ANALYSIS OF DATA SOURCES
- CONCEPTUAL DESIGN
- PROOF-OF-CONCEPT
- PROCESS INTELLIGENCE
- **BUSINESS RECOMMENDATIONS**

BUSINESS RECOMMENDATION

Adjust OLTP:

- Create integrity and business rules (constraints)
- Improve database design

Use anlytical insights:

- Track your material price

Use process insights:

- Improve pain-points (e. g. bottlenecks)

SUPER-X

Thank you for your attention!

SUPER-X

BACK UP