

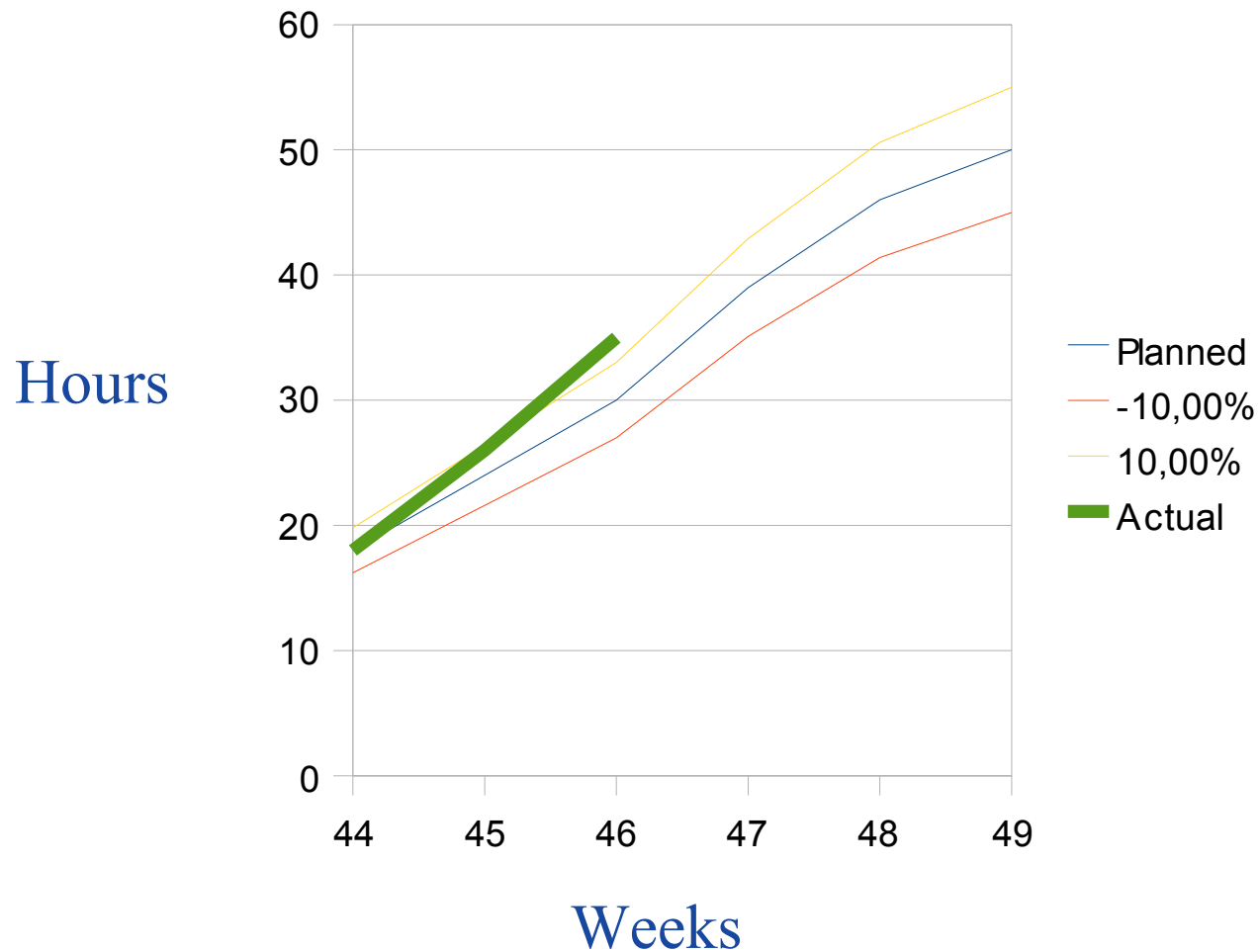
# Data Mining & OLAP Reporting

## Phase: Data Preparation Week 46/2007

**Christoph Johann Szczecina**

**Udi Ibarretxe  
Andoni Albizu  
Dario Casas  
Borja Uriarte**

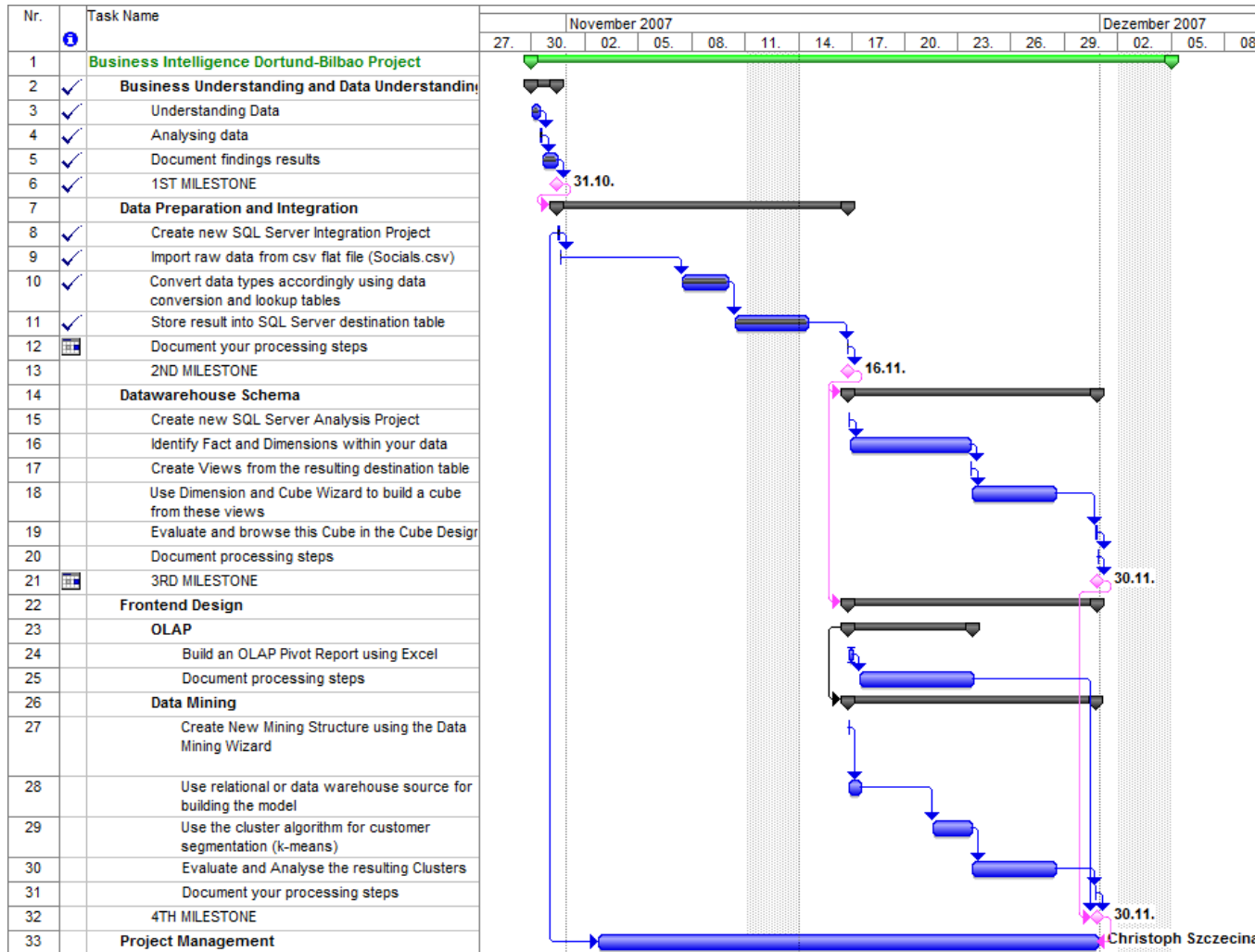
## Budget: Man-Hours



# Progress Overview

Nr.	Task Name	Dauer	Anfang	Ende	% Abgeschlossen
1	<b>Business Intelligence Dortmund-Bilbao Project</b>	<b>50 Std.</b>	<b>Di 30.10.07</b>	<b>Mi 05.12.07</b>	<b>26%</b>
2	✓ <b>Business Understanding and Data Understanding</b>	<b>10,5 Std.</b>	<b>Di 30.10.07</b>	<b>Mi 31.10.07</b>	<b>100%</b>
3	✓ Understanding Data	2 Std.	Di 30.10.07	Di 30.10.07	100%
4	✓ Analysing data	3,5 Std.	Di 30.10.07	Di 30.10.07	100%
5	✓ Document findings results	5 Std.	Di 30.10.07	Mi 31.10.07	100%
6	✓ 1ST MILESTONE	0 Std.	Mi 31.10.07	Mi 31.10.07	100%
7	<b>Data Preparation and Integration</b>	<b>21,5 Std.</b>	<b>Mi 31.10.07</b>	<b>Fr 16.11.07</b>	<b>91%</b>
8	✓ Create new SQL Server Integration Project	2 Std.	Mi 31.10.07	Mi 31.10.07	100%
9	✓ Import raw data from csv flat file (Socials.csv)	3 Std.	Mi 31.10.07	Mi 31.10.07	100%
10	✓ Convert data types accordingly using data conversion and lookup tables	6 Std.	Mi 07.11.07	Fr 09.11.07	100%
11	✓ Store result into SQL Server destination table	2 Std.	Sa 10.11.07	Mi 14.11.07	100%
12	📅 Document your processing steps	2 Std.	Fr 16.11.07	Fr 16.11.07	30%
13	2ND MILESTONE	0 Std.	Fr 16.11.07	Fr 16.11.07	0%
14	<b>Datawarehouse Schema</b>	<b>16 Std.</b>	<b>Fr 16.11.07</b>	<b>Fr 30.11.07</b>	<b>0%</b>
15	Create new SQL Server Analysis Project	2 Std.	Fr 16.11.07	Fr 16.11.07	0%
16	Identify Fact and Dimensions within your data	4 Std.	Fr 16.11.07	Fr 23.11.07	0%
17	Create Views from the resulting destination table	2 Std.	Fr 23.11.07	Fr 23.11.07	0%
18	Use Dimension and Cube Wizard to build a cube from these views	4 Std.	Fr 23.11.07	Mi 28.11.07	0%
19	Evaluate and browse this Cube in the Cube Design	3 Std.	Fr 30.11.07	Fr 30.11.07	0%
20	Document processing steps	1 Std.	Fr 30.11.07	Fr 30.11.07	0%
21	📅 3RD MILESTONE	0 Std.	Fr 30.11.07	Fr 30.11.07	0%
22	<b>Frontend Design</b>	<b>16 Std.</b>	<b>Fr 16.11.07</b>	<b>Fr 30.11.07</b>	<b>0%</b>
23	<b>OLAP</b>	<b>9 Std.</b>	<b>Fr 16.11.07</b>	<b>Fr 23.11.07</b>	<b>0%</b>
24	Build an OLAP Pivot Report using Excel	3 Std.	Fr 16.11.07	Fr 16.11.07	0%
25	Document processing steps	6 Std.	Sa 17.11.07	Fr 23.11.07	0%
26	<b>Data Mining</b>	<b>14 Std.</b>	<b>Fr 16.11.07</b>	<b>Fr 30.11.07</b>	<b>0%</b>
27	Create New Mining Structure using the Data Mining Wizard	1 Std.	Fr 16.11.07	Fr 16.11.07	0%
28	Use relational or data warehouse source for building the model	3 Std.	Fr 16.11.07	Sa 17.11.07	0%
29	Use the cluster algorithm for customer segmentation (k-means)	4 Std.	Mi 21.11.07	Fr 23.11.07	0%
30	Evaluate and Analyse the resulting Clusters	4 Std.	Fr 23.11.07	Mi 28.11.07	0%
31	Document your processing steps	2 Std.	Fr 30.11.07	Fr 30.11.07	0%
32	4TH MILESTONE	0 Std.	Fr 30.11.07	Fr 30.11.07	0%



# Gantt Chart



# Since Last Report: Planned vs. Actual

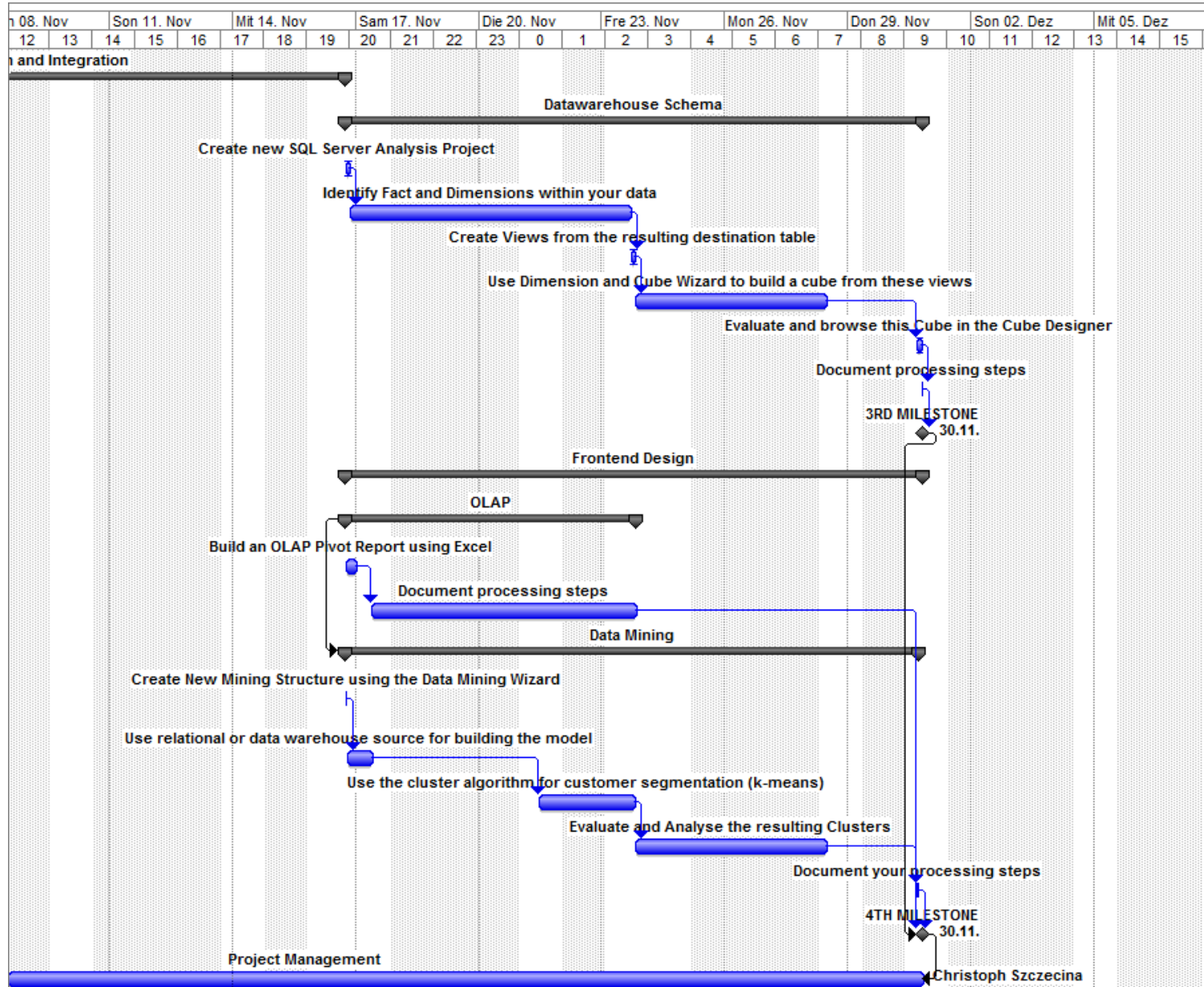
- Current activities
  - Convert data types accordingly using data conversion and lookup tables ●
  - Store result into SQL Server destination table ●

# Current Activities

Nr.		Task Name	Dauer	Anfang	Ende	% abgeschlosse
7		<b>Data Preparation and Integration</b>	<b>21,5 Std.</b>	<b>Mi 31.10.07</b>	<b>Fr 16.11.07</b>	<b>91%</b>
8	<input checked="" type="checkbox"/>	Create new SQL Server Integration Project	2 Std.	Mi 31.10.07	Mi 31.10.07	100%
9	<input checked="" type="checkbox"/>	Import raw data from csv flat file (Socials.csv)	3 Std.	Mi 31.10.07	Mi 31.10.07	100%
10	<input checked="" type="checkbox"/>	Convert data types accordingly using data conversion and lookup tables	6 Std.	Mi 07.11.07	Fr 09.11.07	100%
11	<input checked="" type="checkbox"/>	Store result into SQL Server destination table	2 Std.	Sa 10.11.07	Mi 14.11.07	100%
12		Document your processing steps	2 Std.	Fr 16.11.07	Fr 16.11.07	30%
13		2ND MILESTONE	0 Std.	Fr 16.11.07	Fr 16.11.07	0%

## Until Next Report: Planned

- Documenting steps of Data Preparation
- Create new SQL Server Analysis Project
- Identify Fact and Dimensions within your data
- Build an OLAP Pivot Report using Excel
- Create New Mining Structure using the Data Mining Wizard



## Understanding the Data

	<b>MISSING VALUES %</b>
SEASON	0
YEAR	0
AGENCY	0,03%
DESTINATION COUNTRY	0,31%
DESTINATION REGION	0,32%
INCOME	26,32%
AGE	0,51%
BOOKINGS	0

# Risks

- Wrong time prediction
  - Description: Wrong scheduling, Spanish members only being able to contribute on weekends
  - Action plan: adding resources, project manager
  - Status: High
  - Responsible: whole team
- Intercommunication within the project team
  - Description: Lack of communication
  - Action plan: Regular reporting E-Mails
  - Status: Average
  - Responsible: project leader

## QA Status

- Regular Reporting E-Mails
  - Keep team members on track
  - Double check of results
  - Allows additional improvements through members
- Constant updates to Time Plan
  - Flexible time allocation
- Backups
  - Comparison to former values
  - Security



**thank you!**

**[support@kyynel.biz](mailto:support@kyynel.biz)**