Data Warehouse for Soccer Team
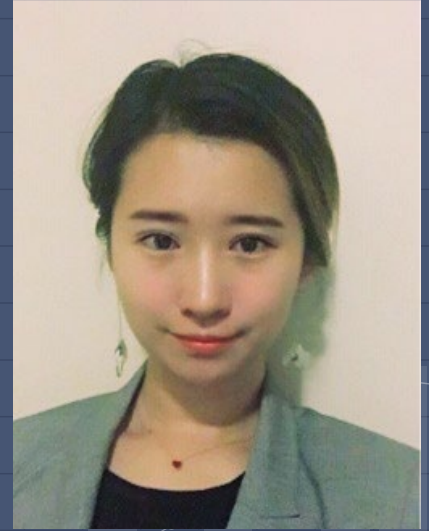
Gaurav Kutemate
Abhijay Sharma
Angel Yang

Abhijay Sharma

Gaurav Kutemate

Angel Yang

Strategic Objective

Having high offensive production has a high impact on fan satisfaction and a small impact on revenue

Approach Description

- Use of Pentaho Ecosystem

- Analyze the impact of various metrics

- Provide organised information

- Support Management to achieve goals

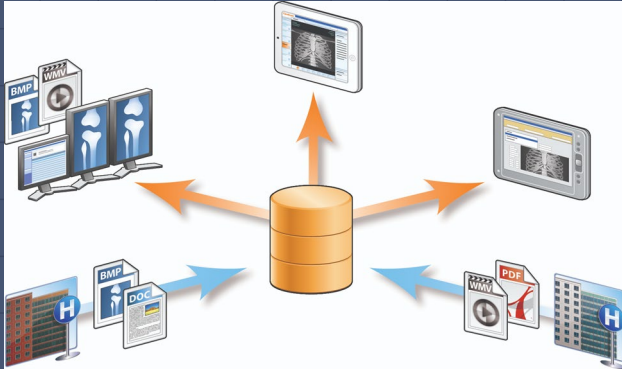# The EPL

- Hotly Contested

- World's best

- Heavy viewership

- Huge fanbase

- Offensive production a key

Why a Data Warehouse?

Need for Centralised storage
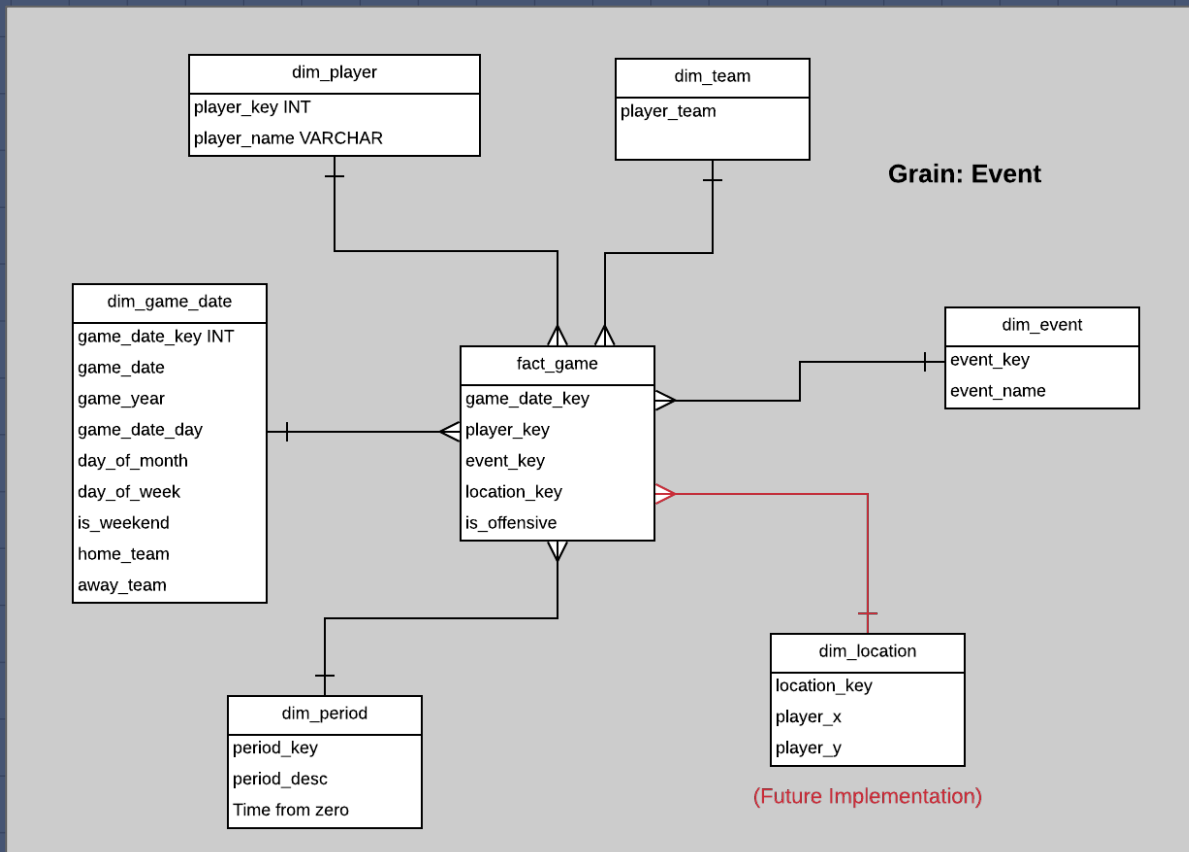
Analytical abilities

Support Soccer operations

Bus Matrix

Business processes	Game Date	Team	Player	Location	Period	Event
Manage player personnel strategy						
Acquire players		✓	✓			
Divest players		✓	✓			
Develop players						
On pitch, game time training	✓	✓	✓			✓
Off pitch Training			✓	✓		✓
Manage injuries						
Physical condition management			✓			
Manage fitness						
Diet management			✓			
Manage player personnel tactics						
Goalie Tactics			✓		✓	
Forward Tactics			✓	✓	✓	
Midfiled Tactics			✓	✓	✓	
Defensive Tactics			✓	✓	✓	
Manage game/opponent tactics						
Formations managements			✓	✓		
Player instructions			✓	✓		
Fouls management	✓					✓
Entertain fans						
Offense productions	✓	✓	✓	✓	✓	✓
Goal productions	✓		✓			✓

Logical Model

Grain: Event

dim_player
- player_key INT
- player_name VARCHAR

dim_team
- player_team

dim_game_date
- game_date_key INT
- game_date
- game_year
- game_date_day
- day_of_month
- day_of_week
- is_weekend
- home_team
- away_team

fact_game
- game_date_key
- player_key
- event_key
- location_key
- is_offensive

dim_event
- event_key
- event_name

dim_period
- period_key
- period_desc
- Time from zero

dim_location
- location_key
- player_x
- player_y

(Future Implementation)

What have we accomplished?

- Created Dimensional Model

- Conceptualized strategic objectives

- Built Player analysis

- Team selection procedure

What needs refinement?

- Injury management

- Team utilization

Planned for future

- Combine event and location

- Zoned categorization
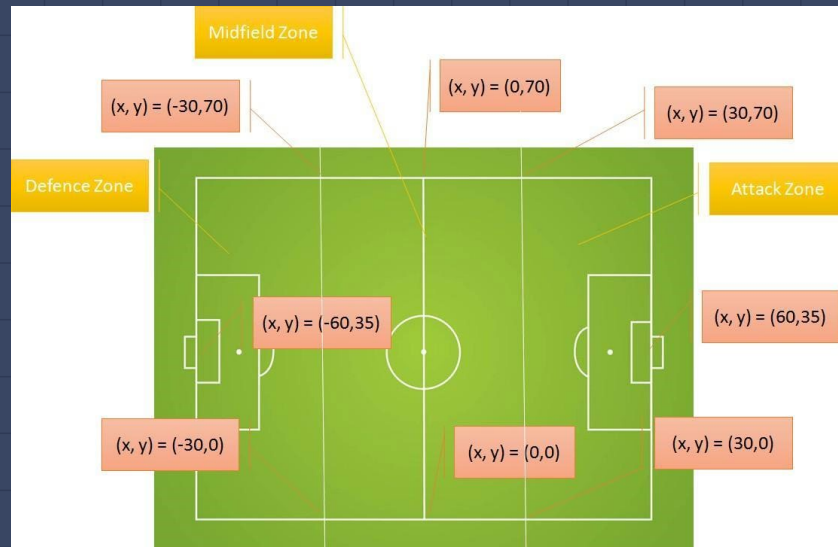
Physical Design

- Creating tables using MySQL

- ETL in Pentaho Data Integration

- OLAP cubes in Schema Workbench

- Reports and Analyses

Creating SQL tables

ETL Transformations–Dimensions
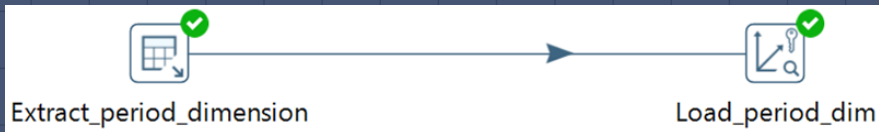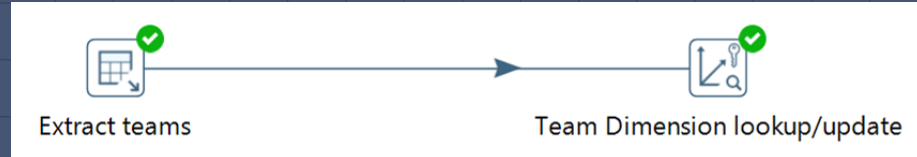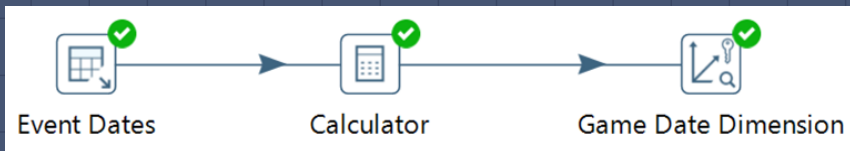
ETL Transformations—Fact Table

OLAP Cubes

Reports

July 13, 2019 @ 11:15

Report on Offensive Events of each team

Teams	Period	Number of offensive events
team10	First Half	127
team10	Second Half	171
team13	First Half	83
team13	Second Half	66
team15	First Half	137
team15	Second Half	118
team16	First Half	190
team16	Second Half	200
team17	First Half	61
team17	Second Half	86
team18	First Half	132
team18	Second Half	122
team19	First Half	61
team19	Second Half	103
team2	First Half	105
team2	Second Half	123
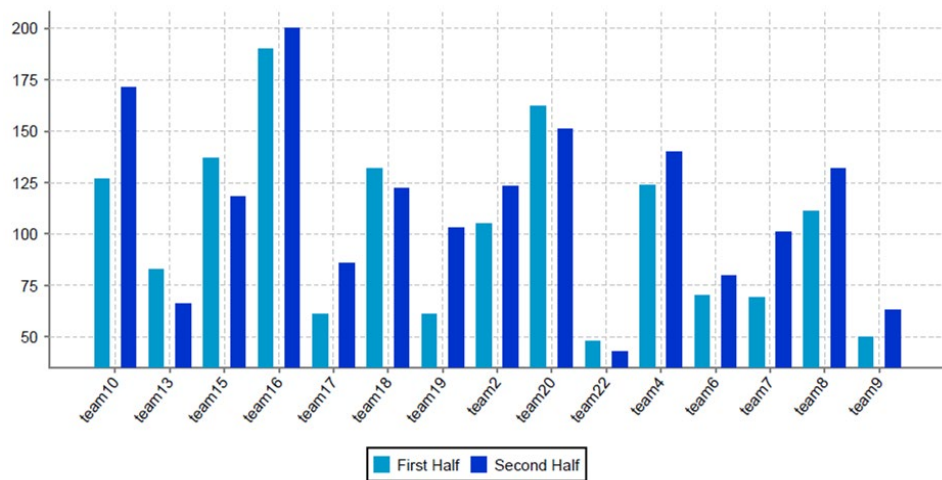
Conclusion….

- Small but a crucial step towards creating an analytics environment
- Capable of being customized
- Serves different processes

Challenges

- Conceptualizing the project approach
- Designing the Bus matrix
- EDA of data source due to inconsistencies
- Finalizing of Dimensions and grain
- Navigating on Pentaho Report Designer

Questions?