

CIS 9440 - Data Warehousing and Analytics

Class #9





This week's Office Hours

Changed from Thursday to **Friday at 11:00am**



HW #2 Comments

I received many questions within the last 2 days

- Will provide some help and move the due date back to Friday
 - The number of attempts you may submit on Blackboard was updated from 1 to 2
 - Must submit by Friday at midnight



HW #2 Help

CSV is many columns:

https://drive.google.com/file/d/1adkkH5xlRLj_y78KdITb1sRVuKbsZJc9/view

- Must be transposed into 3 columns: Country, Year, and Income per Person
 - Use excel
 - Or, use pandas melt function: [link1](#)
 - `data.melt(['country'], var_name = 'Year')`

Week 9 Class Overview:

1. Warmup SQL
2. Downloading Tableau for next week
3. ETL Workshop #3
4. Intro: BI Application Design and Development

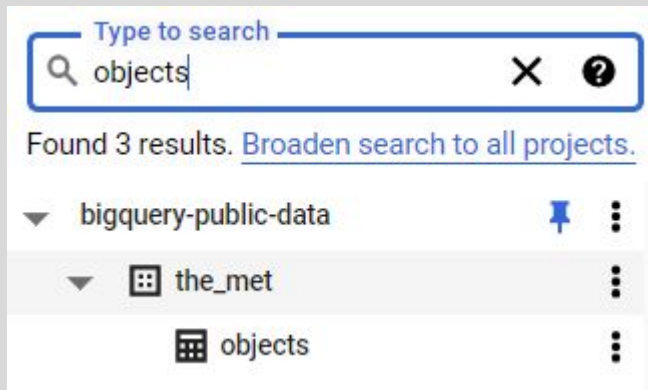
Week 9 Class Overview:

- 1. Warmup SQL**
2. Downloading Tableau for next week
3. ETL Workshop #3
4. Intro: BI Application Design and Development



Which department has the most public objects at the Met?

Use the following BigQuery table:



Week 9 Class Overview:

1. Warmup SQL
- 2. Downloading Tableau for next week**
3. ETL Workshop #3
4. Intro: BI Application Design and Development



Downloading Tableau

1. Poll
2. <https://www.tableau.com/academic/students>

Week 9 Class Overview:

1. Warmup SQL
2. Downloading Tableau for next week
- 3. ETL Workshop #3**
4. Intro: BI Application Design and Development



ETL Workshop #3 outline

1. Review our Project Goal
2. Look at our datasets: [first](#), [second](#)
3. Review the Dimensional Model
4. Complete ETL process together
5. Create your own KPI's
6. On your own (individual or in a group): Answer your Project Goal with the KPI's



Google Form for Steps 5 and 6

- <https://forms.gle/anj8jHZ2n2z9jhAj6>



Project Goal (challenge)

- Identify the best locations in NYC to apply for a Sidewalk Cafe (outdoor seating).



Project datasets

- Look at our datasets: [first](#), [second](#)

Dimensional Model

Dimension	
Name: location_dim	
location_id (PK)	
latitude	
longitude	
street	
zip_code	
borough	

Fact Table	
sidewalk_fact	
location_id (FK)	
business_id (FK)	
square_feet	
tables	
chairs	

Dimension	
Name: business_dim	
business_id (PK)	
company_name	
business_name	
sidewalk_cafe_type	
sidewalk_license_status	



ETL Process

- Download the Jupyter Notebook “Sidewalk_ETL” from Blackboard



Create

1. Review our Project Goal
2. Look at our datasets: [first](#), [second](#)
3. Review the Dimensional Model
4. Complete ETL process together
5. Create your own KPI's
6. On your own (individual or in a group): Answer your Project Goal with the KPI's



Create your own KPI's

- KPI #1:
- KPI #2:
- KPI #3:



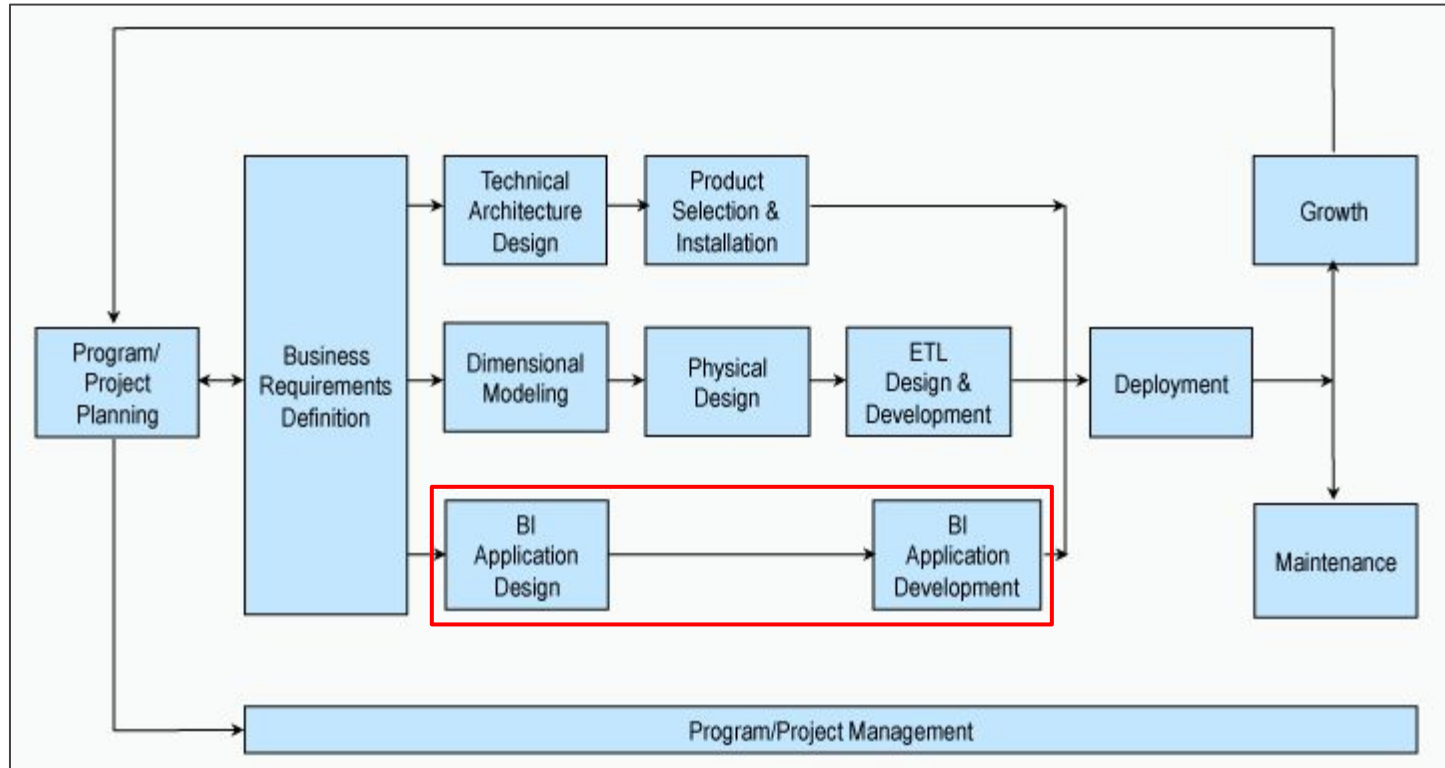
Conclusion

- On your own (individual or in a group): Answer your Project Goal with the KPI's in the Google Form

Week 9 Class Overview:

1. Warmup SQL
2. Downloading Tableau for next week
3. ETL Workshop #3
- 4. Intro: BI Application Design and Development**

Kimball Lifecycle - Business Intelligence (BI)



Notice, you cannot start “BI Application Development” (actually building your BI Applications) until both the Middle and Top Kimball Lifecycle tracks are complete.



What is Business Intelligence?

Business Intelligence (BI) is leveraging or analyzing data to provide transparent, strategic, and actionable insights to an organization.

Business Intelligence requires 2 components:

1. **Transformed Data**, for data-driven decision making
2. Business Intelligence Platform



What is Business Intelligence?

Business Intelligence (BI) is leveraging or analyzing data to provide transparent, strategic, and actionable insights to an organization.

Business Intelligence requires 2 components:

1. **Transformed Data**, for data-driven decision making
2. Business Intelligence Platform



Business Intelligence Platform

A business intelligence (BI) platform is **technology** that helps businesses gather, understand, and visualize their data. It serves as the backbone of a company's business intelligence strategy, which is how a company uses information to make better decisions.

When implemented well, modern business intelligence platforms **make your data usable**.



Business Intelligence Platform Examples

- [Microsoft Power BI](#)
- [Tableau](#)
 - Tableau [Viz of the Day](#)
 - Tableau yearly conference
- [Looker](#)
 - GCP Integration
- [Qlik Sense](#)
- [ThoughtSpot](#)
 - “Life without dashboards”
- [Oracle Analytics](#)
 - “Enterprise” marketed

Poll! Favorite BI Platform



How are BI Platforms ranked? Who is Gartner?

Gartner, Inc, officially known as Gartner, is a global research and advisory firm providing information, advice, and tools for leaders in IT, finance, HR, and many other verticals.

Gartner publishes the **Gartner Magic Quadrant for BI Platforms** every year, and it is taken very seriously in the Data Warehousing space.



Gartner BI Platforms Ratings

- Varying costs for each Platform
- Each Platform has strength and weaknesses; best option depends on your project

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



Source: Gartner (March 2022)

2020



2022

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



Source: Gartner (March 2022)

2016



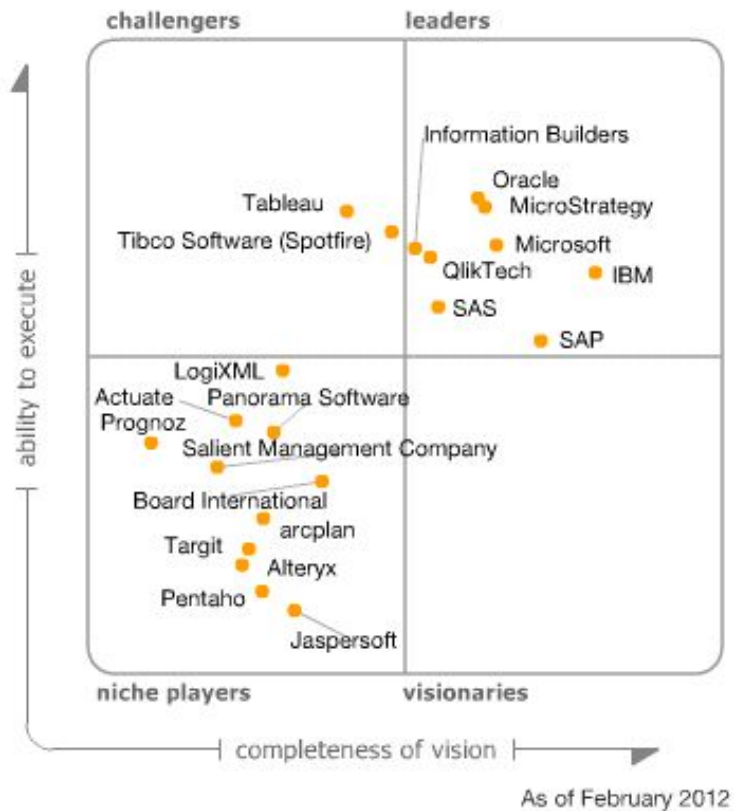
2022

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



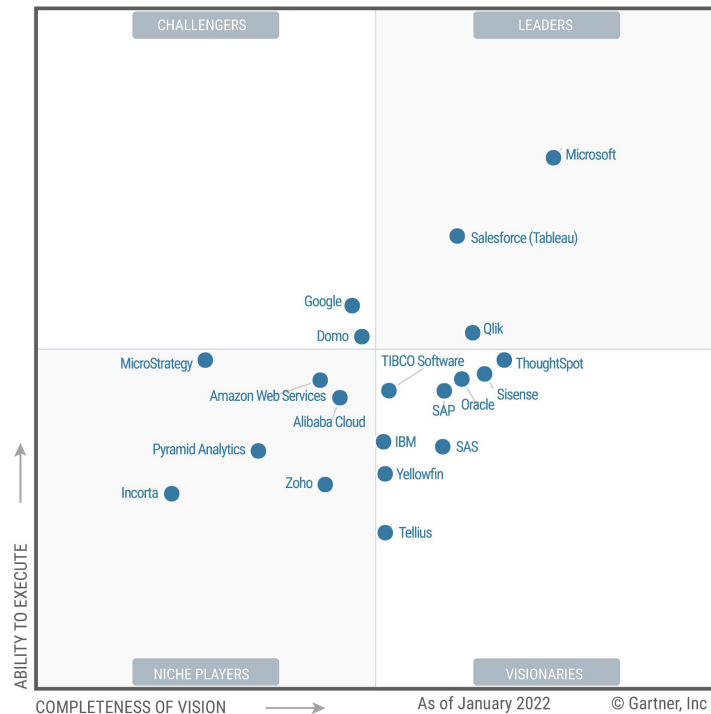
Source: Gartner (March 2022)

2012



2022

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



Source: Gartner (March 2022)



BI Applications, what are the differences?

A BI Application is a BI Platform deliverable that is developed to for specific users to measure a specific KPI.

Examples:

- Dashboard
- Scorecard
- Report
- Data Visualization
- Ad hoc query
- Predictive Model



How does Business Intelligence relate to Analytics?

- What is Analytics?
 - systematic computational analysis of data
- What is Business Intelligence?
- Business Intelligence is a vehicle to deliver analytics.



When to use a Report?

- broad, historic, and **static** data.
 - Reports are not very flexible.
- recurring basis: daily, weekly, monthly, or quarterly.
- wide audiences - like an entire division of an organization
 - Reports may be drilled in on different data by business unit
- data in a table format



When to use a Dashboard?

- **multiple KPI's** in a singular view.
- **data that will be interacted with**, drilled upon, or manipulated by the users.
- near-live data, and used on a more frequent basis than a Report.
- unlike reports, Dashboards do not quickly deliver a message. They allow the user to find the most important messages.



When to use a Scorecard?

- measuring performance against a benchmark.
 - Example: Current Sales vs Quarterly Sales Target
- more specific than reports, with static views of very specific KPI's
- Scorecards may be for wide or narrow audiences.
- Like reports, Scorecards are not very flexible.



When to use a Data Visualization?

- generated ad hoc, for specific presentations, emails, or research.
- One place you may find a Data Visualization is a Team presentation to the company.
 - targeted purpose and message
 - Oftentimes, a Data Visualization is created from a Dashboard. In this case, it's a specific cut from a Dashboards full data.



Poll!

Which of the following are BI Applications?





Homework:

1. HW #2 due this Friday
 2. Final Project Milestone #3 on Blackboard, due 11/9/22
- 