ReDI School of
Digital Integration

Business Perspective

Lecture 5

We use tech to connect human potential and opportunity with dignity & humility

Today's topics

- Different types of analytics
- Ethical considerations
- Challenges with data
- Data roles
- Describe visuals

Break

- Data querying
- Data visualization

Different Analytics

Descriptive Analytics

Looks at data to examine, understand, and describe something that **has already happened**.

Looks at events in the past → Identify specific patterns
Commonly used visuals: pie charts, bar charts, tables, line graphs

Example:
How many sales occurred in the last quarter?

Diagnostic Analytics

Provides a deeper analytics by seeking to understand the “**why**” behind what happened.

You must first understand what happened to be able to analyze what will happen.

Example:
Why did sales drop in January of last year?

Predictive Analytics

Relies on historical data, past trends, and assumptions to answer questions about what **is likely to happen** in the future.

Techniques:
Regression analysis, forecasting, multivariate statistics, predictive modeling

Requirements for these techniques are large amounts of high-quality and historic data. They are also done using programming such as R and Python.

Example:
How will the next marketing campaign impact the customer engagement?

Prescriptive Analytics

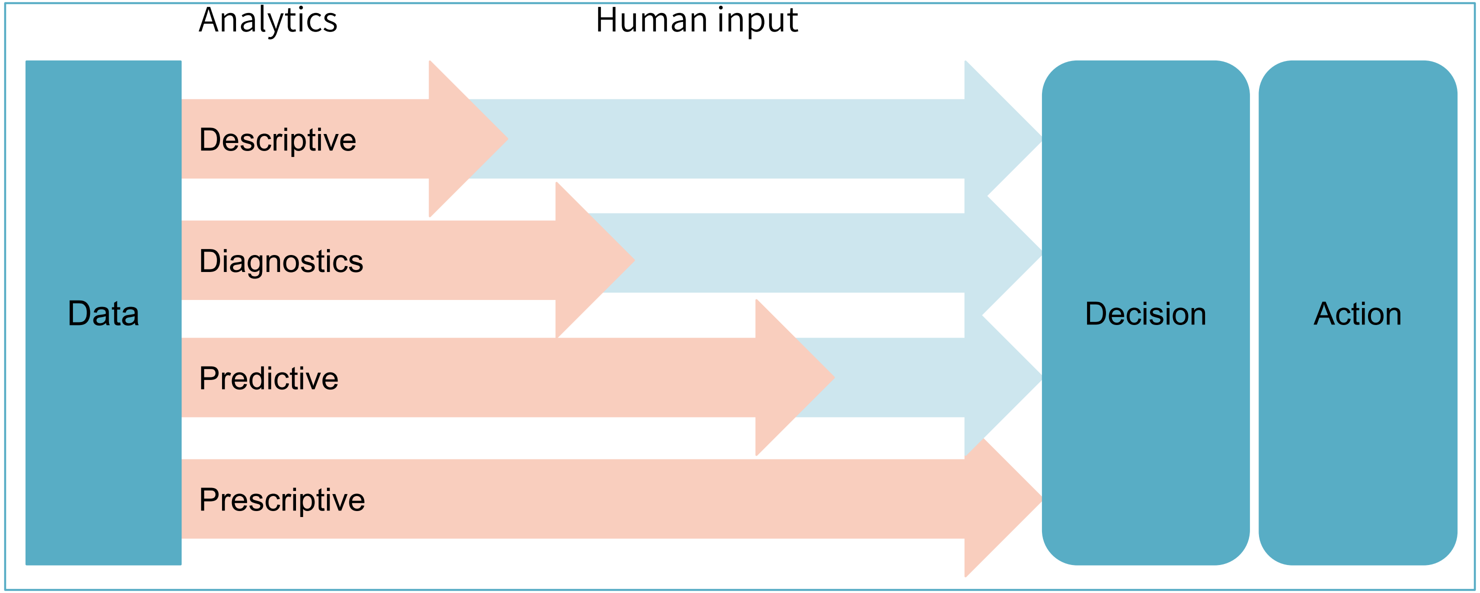
Identifies specific actions an individual or organization should take to **reach future targets or goals**.

Techniques:
Graph analysis, simulation, neural networks, machine learning

This is one of the most advanced level of analytics and it heavily depends on the accuracy of the descriptive, diagnostic, and predictive analysis. It also heavily relies on quality of data and the appropriate data architecture and expertise.

Organization will be able to **make decisions based on highly analyzed facts** rather than instinct.

Analytics in decision making



Cognitive Analytics

Cognitive analytics is an advanced type of analytics that uses **artificial intelligence and machine learning** to identify patterns and trends in large sets of data.

It can be used to uncover insights that may not have been visible before, and to create predictive models that can help organizations make better decisions. Organizations are using cognitive analytics to tap into unstructured data sources such as images, emails, text documents, and social posts.

Examples:
Apple's Siri, Samsung's Bixby

Ethical considerations

Data Collection

- **Informed consent**

All subjects have to be *clearly informed* on how the data is to be used. Based on this information they have to clearly *give consent.*

- **Collection bias**

Bias at the collection stage means that the data gathered is *not truly representative* of the situation you are trying to investigate.

- **Personally identifiable information (PII)**

Have to find ways to *minimize the exposure* of personally identifiable information for example through *anonymization*. Always keep *GDPR* in mind!

Data Storage

- **Data security**

A plan to protect and secure data, such as encryption, access controls, etc.

- **Right to be forgotten**

The subjects have to have the right request their personal information to be removed

- **Data retention plan**

Does the data have to be deleted after it is no longer needed?

Analysis

- **Missing perspective**

Could the analysis have missing perspective? How can we address this?

- **Privacy**

Are PII sensitive data have been removed/addressed in the data before analysing/visualizing?

- **Auditability**

Is the analysis well-documented and can be reproduced later if needed?

Challenges with data

Meaningful data

Companies are now data-driven = lots of data → hard to find insights in a lots of data.

It's impossible to sort and analyze all the data in real-time, which might fail to provide accurate and relevant reports. To make it easy, using appropriate data analytics tools are important.

Multiple sources

The various sources are not necessary the same format or they are hard to join
Employees may not know that the data they need are not in the same source

Scaling

With the rapidly increasing data volume, businesses face the challenge of scaling data analysis. Also, one might find that the software tool is not scalable.

Data culture

Most companies are not data-driven and employees do not know much of DA.

- Inaccurate data (usually manual/human errors)
- Often changes in one system don’t reflect in another system

Data security

Increase in data volume also increases data storage, and that calls for security measures that minimize cyber risks.

Data roles

Data Analyst

“Data Analyst analyzes numeric data and uses it to help companies make better decisions.”

Responsibilities

Pre-processing and data gathering

Representing data via reporting and visualization

Statistical analysis and data interpretation

Data maintenance

Communication with stakeholders

Data Scientist

"A data scientist analyzes and interpret complex data. They are data wranglers who organize (big) data."

Responsibilities

Data analytics and optimization using machine learning and deep learning

Strategic planning for data analytics

Statistical analysis and data interpretation

Data Engineer

"Data Engineer involves in preparing data. They develop, constructs, tests & maintain complete architecture."

Responsibilities

Develop, test, and maintain architectures

Deploy machine learning and statistical models

Building pipelines

Creating ETL/ELT operations

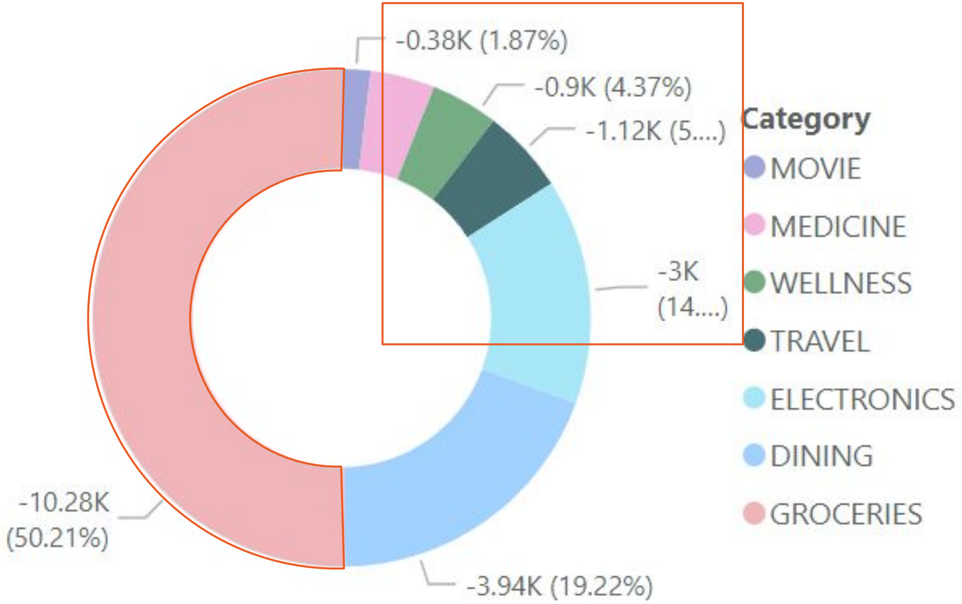
Describe Visuals


Total Defect Qty by Month and Year
6M
4M
2M
0M
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Year
2013
2014

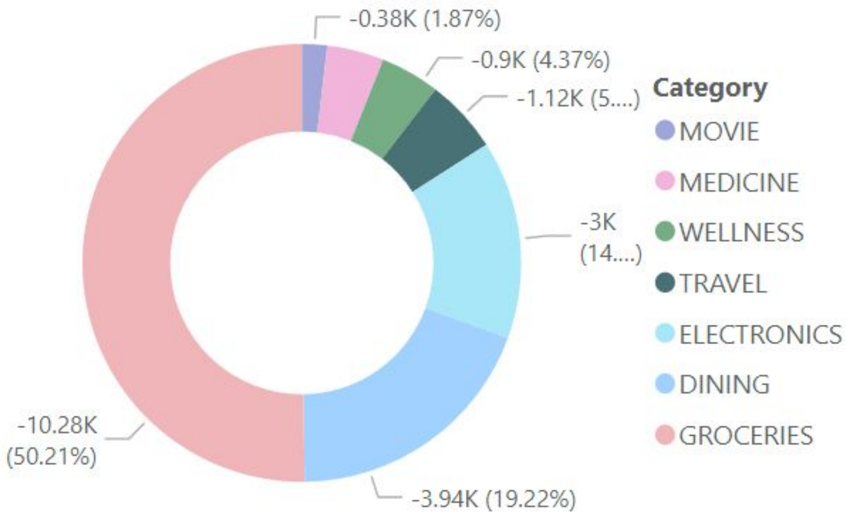
In 2013, April and June were the months with the highest amount of defect products. However, 2014 saw an increase in defective products overall, with December being the month with the most.

There were also three noteworthy drops in the amount of defect products in both years: between June and August in 2013 and between January and February, as well as between October and December in 2014. [Here, if you know the underlying reason, you can explain it]

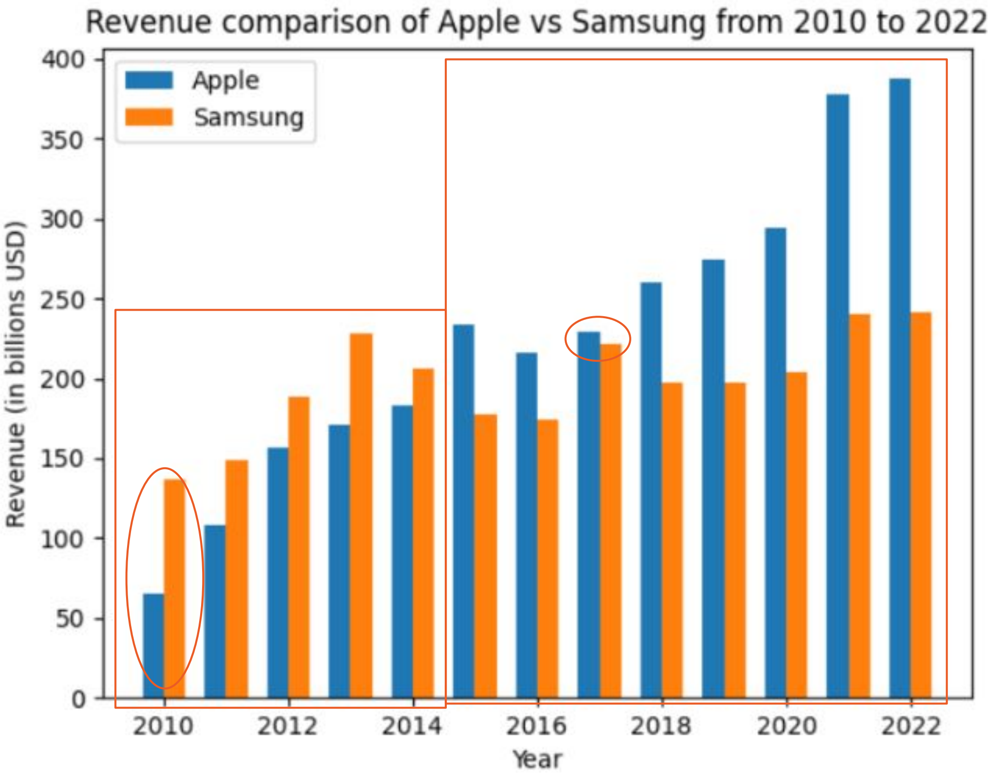
Spendings by category

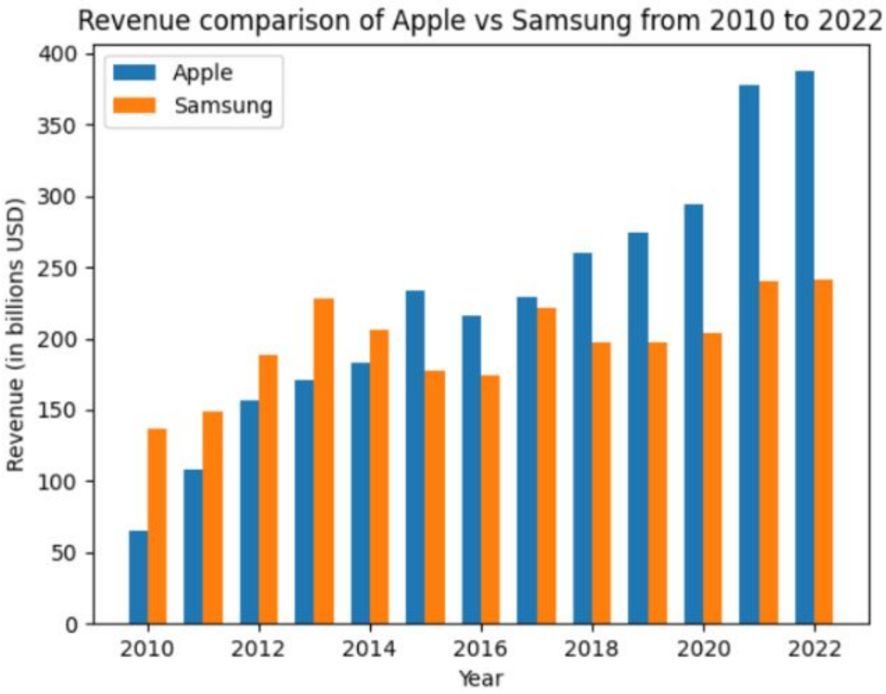


Spendings by category



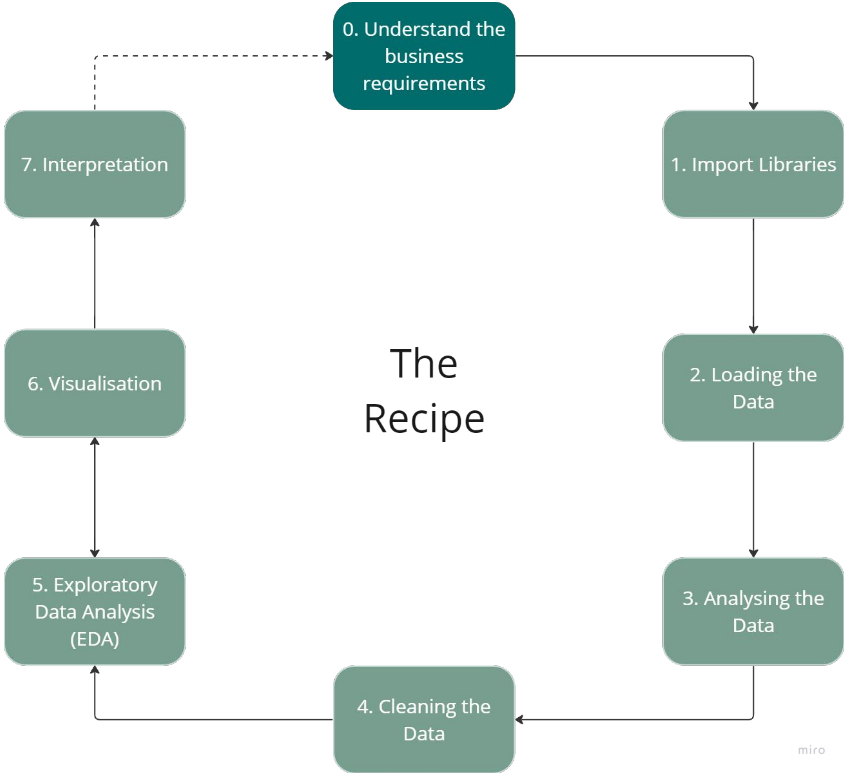
We can see that over half (50.21%) of this family's income is spent on Groceries in a quarter, followed by Dining with almost 20% and Electronics with a lower percentage. To maximize their savings, we suggest reducing spending in these three categories.


Revenue comparison of Apple vs Samsung from 2010 to 2022
Apple
Samsung
Revenue (in billions USD)
0
50
100
150
200
250
300
350
400
2010
2012
2014
2016
2018
2020
2022
Year



This visuals shows the revenue of Samsung and Apple between 2010 and 2022. We can see that from the beginning of the decade until 2014 Samsung's revenue was higher than that of Apple, which later on changed to the benefit of Apple.
In 2010, Samsung's revenue was double of Apples and the closest their revenue has ever been was in 2017.

The recipe for data analysis



After the break, it is exercise time!