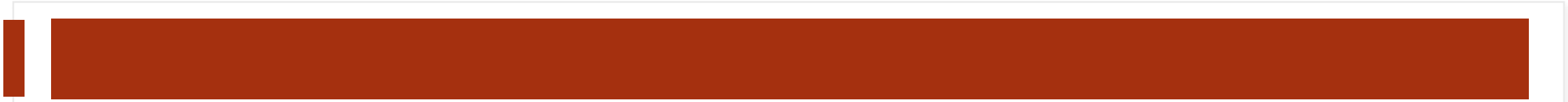
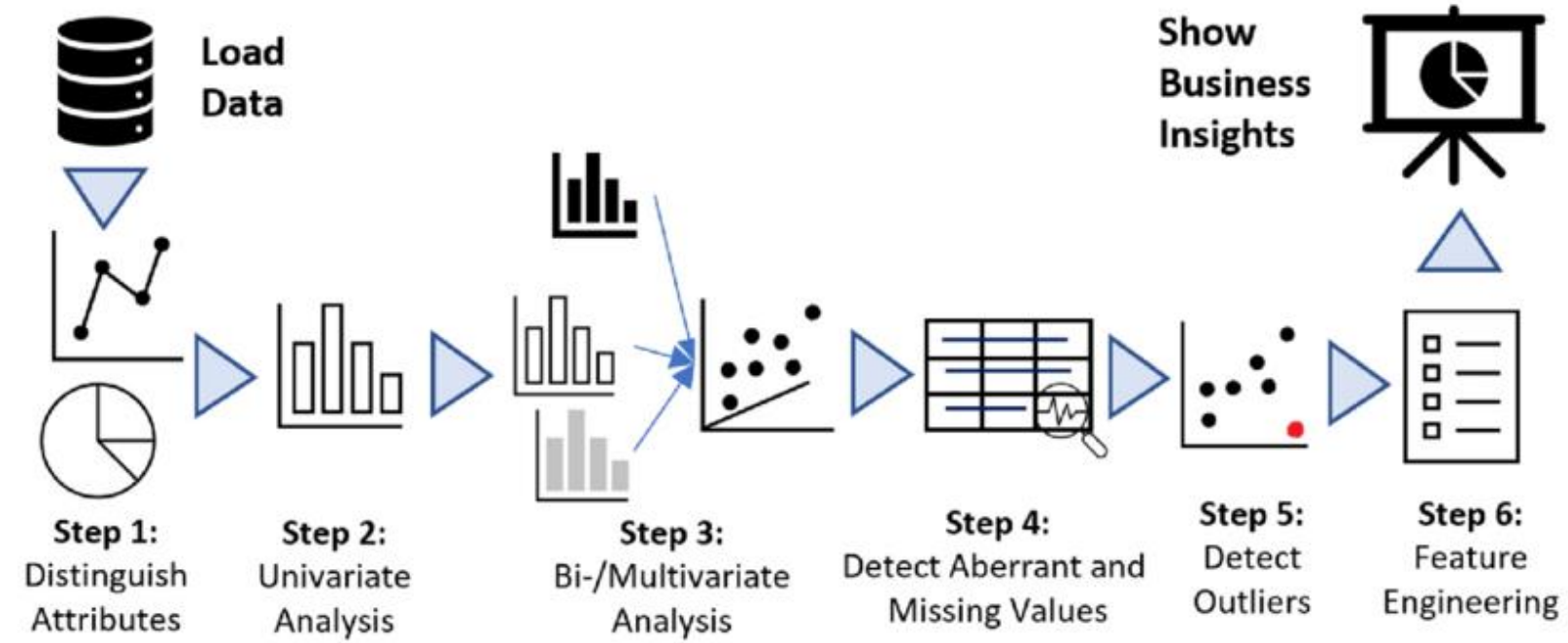
# Welcome Back!

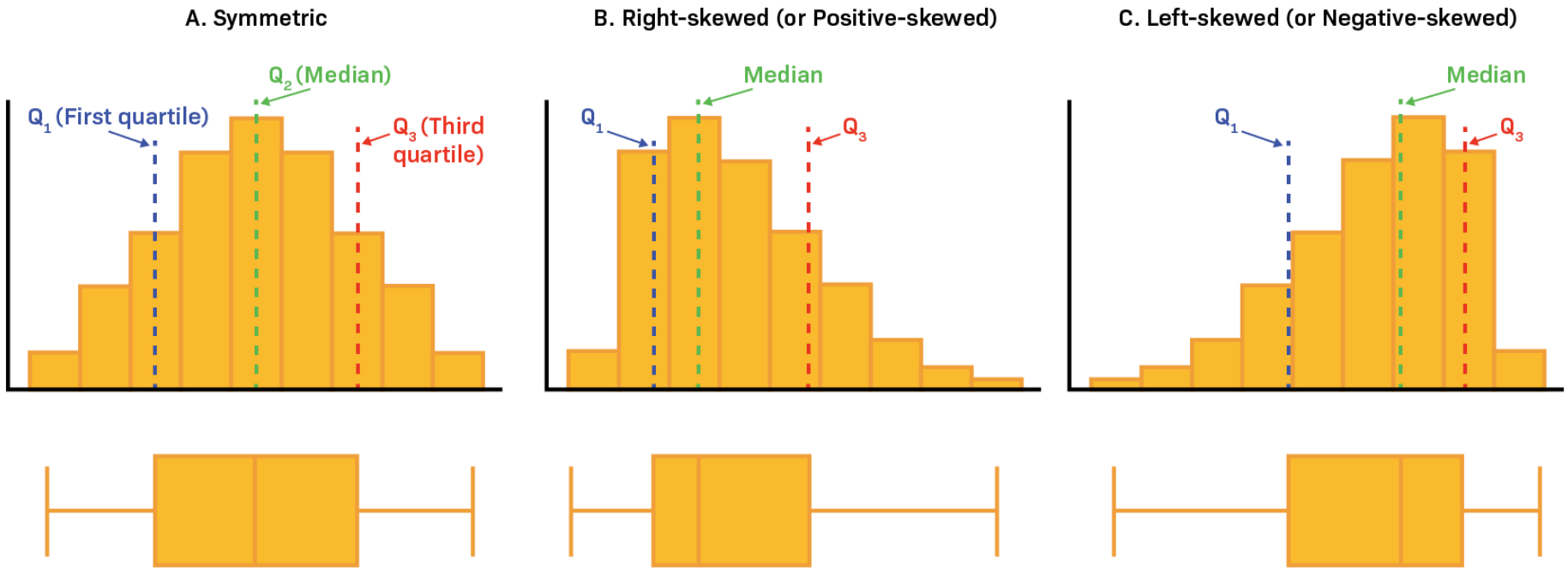


# Recap – Data Wrangling

# Recap - EDA


Load Data
Show Business Insights
Step 1: Distinguish Attributes
Step 2: Univariate Analysis
Step 3: Bi-/Multivariate Analysis
Step 4: Detect Aberrant and Missing Values
Step 5: Detect Outliers
Step 6: Feature Engineering

# Recap – Univariate Analysis



# Recap – Bi/Multivariate Analysis

- Correlation Testing
- Regression (Linear/Logistic)
- ANOVA – Analysis of Variance
- Tukey Test
- T-test
- Chi-square Test
- Normality Testing
  (Shapiro-Wilk, K^2, Kolmogrov-Smirinov, Anderson Darling)

# Identifying the right statistical test

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | ANOVA<br>TUKEY |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | Correlation |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | Chi Squared |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | ShapiroWilk<br>Kolmogrov Smirnov<br>K-2 |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | Linear Regression |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | Independent samples t-test |

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>TUKEY |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | Correlation |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | Chi Squared |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | ShapiroWilk<br>Kolmogrov Smirnov<br>K-2 |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | Linear Regression |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | Independent samples t-test |

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>**TUKEY** |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | Correlation |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | Chi Squared |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | ShapiroWilk<br>Kolmogrov Smirnov<br>K-2 |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | Linear Regression |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | Independent samples t-test |

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>**TUKEY** |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | **Correlation** |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | Chi Squared |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | ShapiroWilk<br>Kolmogrov Smirnov<br>K-2 |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | Linear Regression |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | Independent samples t-test |

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>**TUKEY** |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | **Correlation** |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | **Chi Square** |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | ShapiroWilk<br>Kolmogrov Smirnov<br>K-2 |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | Linear Regression |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | Independent samples t-test |

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>**TUKEY** |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | **Correlation** |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | **Chi Square** |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | **ShapiroWilk**<br>**Kolmogrov Smirinov**<br>**K-2** |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | Linear Regression |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | Independent samples t-test |

**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>**TUKEY** |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | **Correlation** |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | **Chi Square** |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | **ShapiroWilk**<br>**Kolmogrov Smirinov**<br>**K-2** |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | **Linear Regression** |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | **Independent samples t-test** |

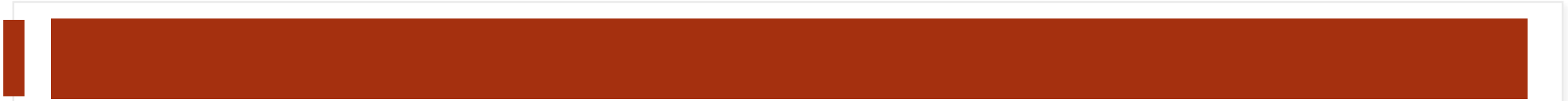
**Identify which statistical test (s) would be best in the following scenarios.**

| | | |
|---|---|---|
| **25** | Comparing the mean test scores of students from CS, BBA or SSLA. If the mean difference is evident, investigate further. | **ANOVA**<br>**TUKEY** |
| **26** | Identifying the strength and direction of the relationship between height and weight for a health profile dataset. | **Correlation** |
| **27** | Analyzing whether job role (Developer, Analyst, Manager) influences preferred working hours (morning, afternoon, night) | **Chi Square** |
| **28** | Checking if the distribution of monthly sales follows a normal pattern | **ShapiroWilk**<br>**Kolmogrov Smirinov**<br>**K-2** |
| **29** | Evaluating the effect of age, years of experience, and certification count on salary | **Linear Regression** |
| **30** | Comparing the average daily sales of two stores (Store A and Store B) to see if there's a significant difference in their performance | **Independent samples T-test** |

# Solving BI Problems

CS 459 Business Intelligence

# What is
# Design Thinking



# DESIGN THINKING

More doing than thinking

ENGINEERING THINKING

Solve your way forward

BUSINESS THINKING

Optimize your way forward

RESEARCH THINKING

Analyze your way forward

DESIGN THINKING

Build your way forward

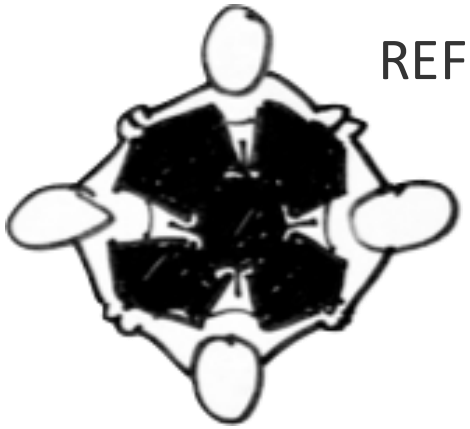
# DESIGN THINKING

Key mindsets and practices

REFRAMING

MINDFULNESS OF PROCESS

STORYTELL



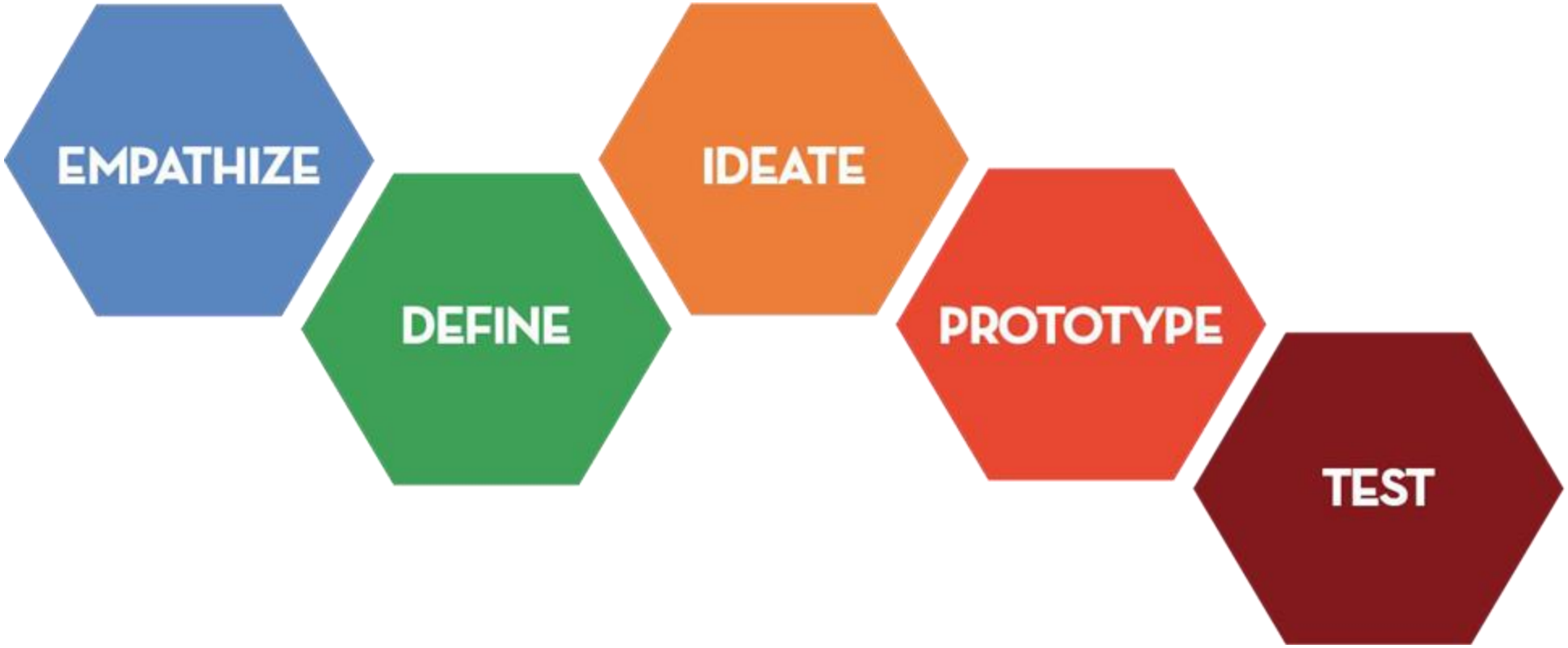
RADICAL COLLABORATION

CURIOSITY

BIAS TOWARD ACTION

# DESIGN THINKING

## The framework



# Empathize
## Understand Stakeholder Needs

Gather deep insights into business problems from stakeholders.

# Define
## Frame the BI Problem

Clearly articulate the BI challenge.

# Ideate
## Explore BI Solutions Creatively

Brainstorm potential BI approaches

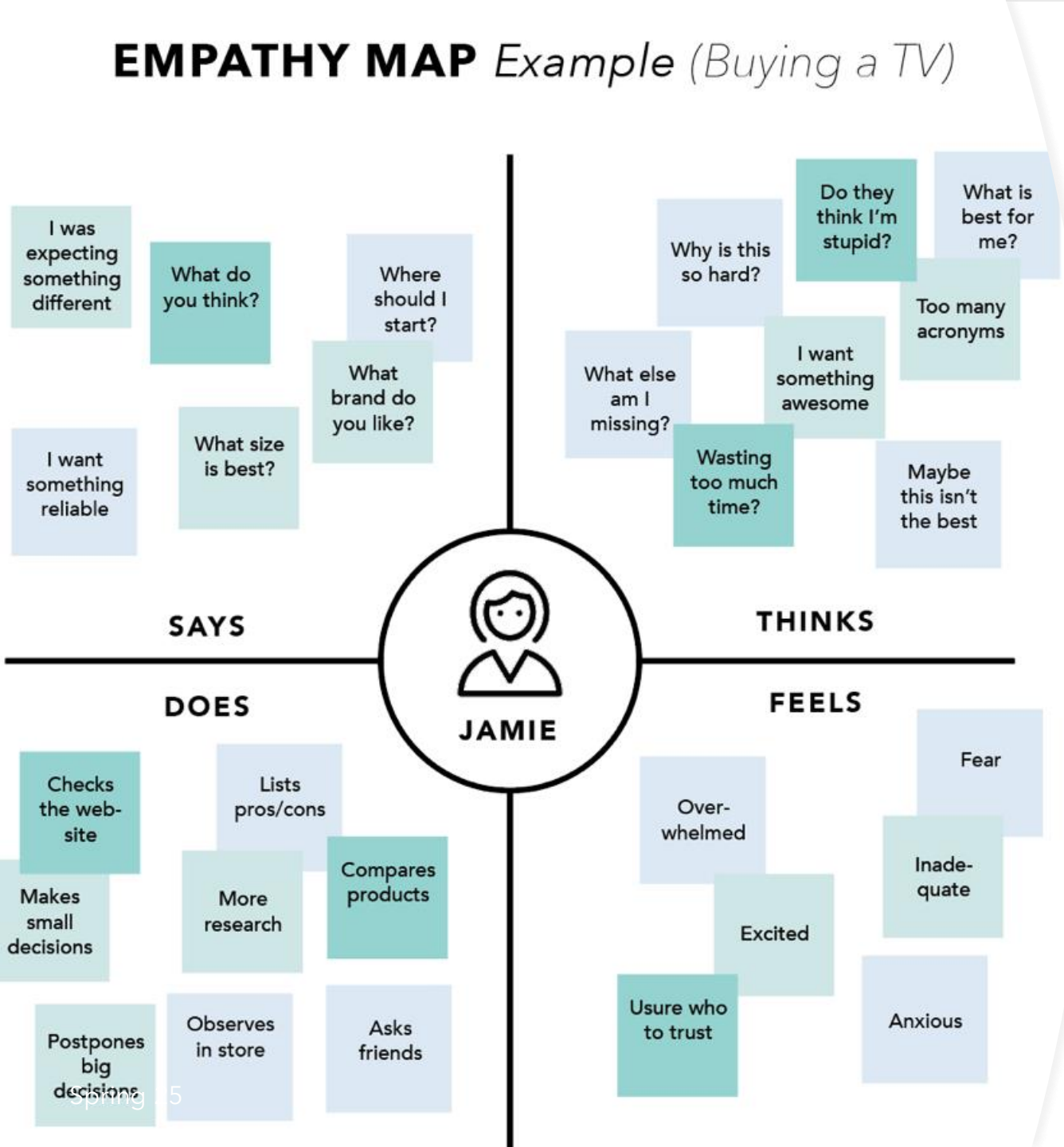
# Prototype
## Build BI Solutions Iteratively

Develop tangible BI deliverables quickly.

# Test
# Validate with Stakeholders

Refine BI solutions based on feedback.

# Empathy Mapping




# Use Empathy Mapping

This tool can help you to understand the thoughts, feelings, and needs of the people who are affected by the problem.

# The Task

- Shuffle around and pair up.
- Consider some problem/goal that you would like to work on.
- Interview the other person around their identified problem.
- Not allowed to ask what do you think/say/do/feel
- Infer from their story
- *5 mins:* Person A interviews Person B and fills out the empathy map
- *5 mins:* Person B interviews Person A and fills out the empathy map

# Examples

- I am working towards getting my dream internship.
- Buying a new __?__
- My GPA is falling
- Graduation is in 3 months, and I have no job offers
- Social media drains my energy
- My business is not doing well – Sales dropped 30% last month