
# DA2I
# Data Storytelling and Dashboard Design

Leandro Nunes de Castro, Ph.D.

ldecastrosilva@fgcu.edu

# Data Storytelling and Dashboard Design

## Summary

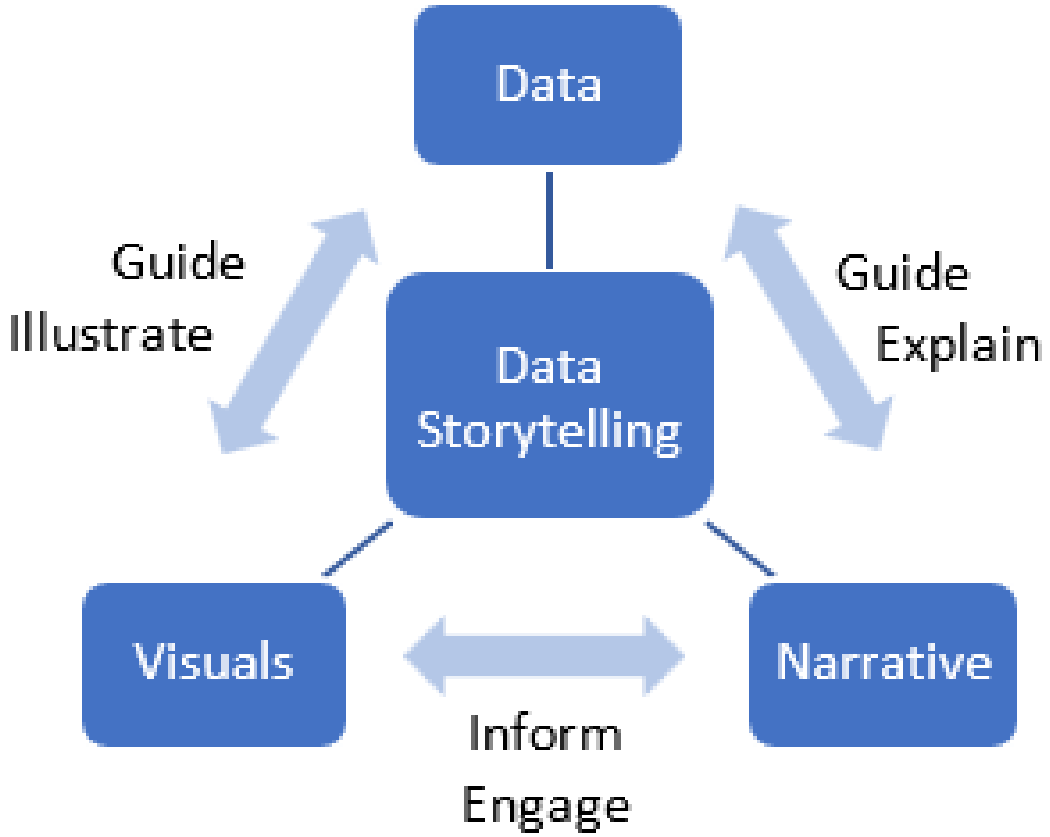
- Data Storytelling
- Dashboard Design
- Case Study

- *Data storytelling* is the practice of building a narrative around data and its accompanying analysis and visualizations.
- Whilst data storytelling is the narrative, *dashboard design* is the process by which the end result of the exploratory data analysis is built and presented in an intuitive and aesthetically pleasing manner.
- Dashboards serve as the visual interface that transform the data analysis process into dynamic and intuitive displays.

# Data Storytelling

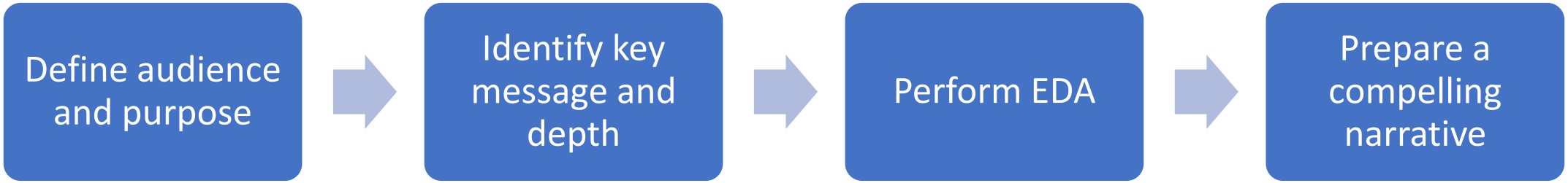
# Storytelling Design Process

Data storytelling involves an effective combination of *data*, *narrative*, and *visuals.*



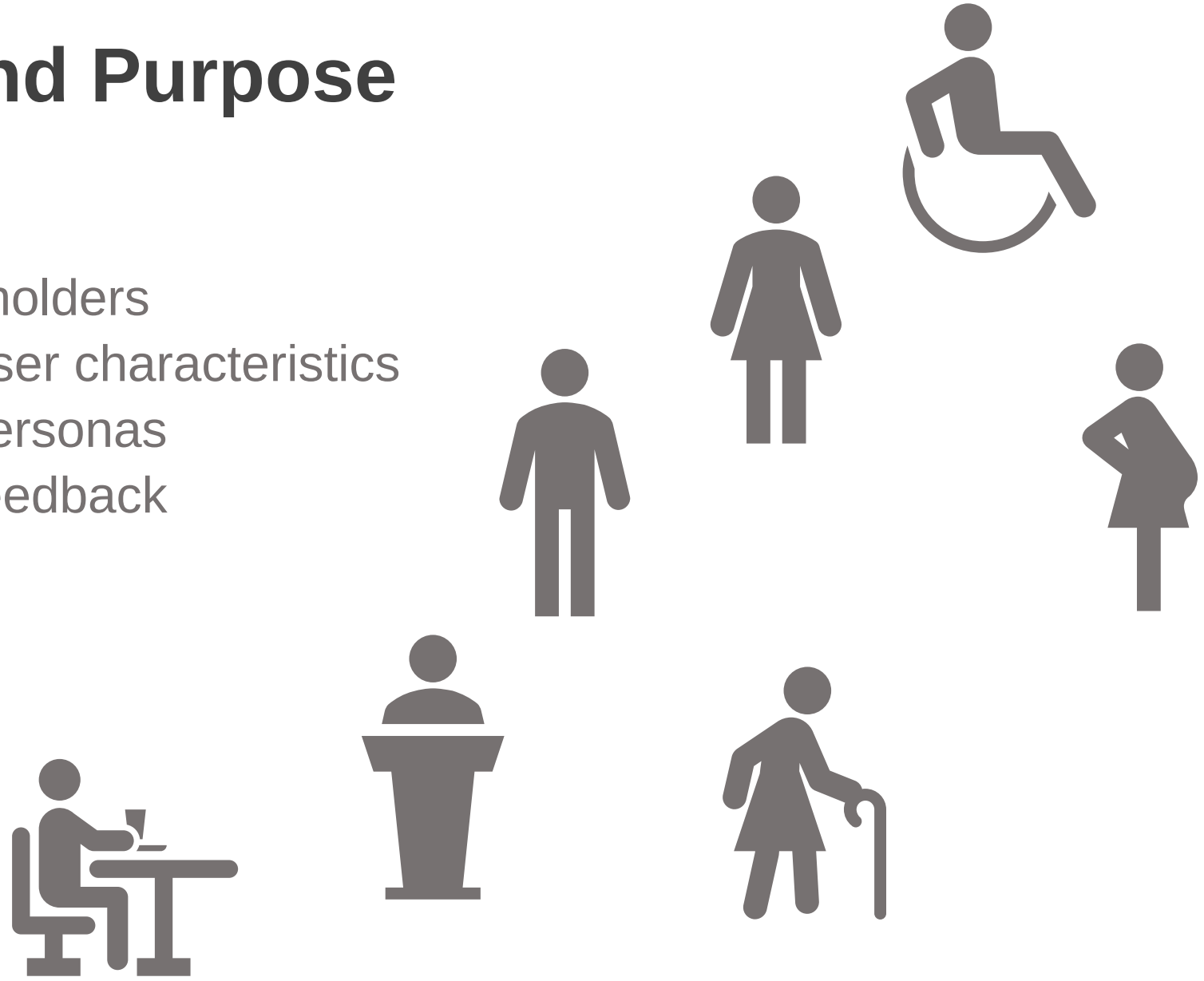
# Storytelling Design Process

- What questions do you want to answer with the data?
- What kind of relationships exist in the data?
- What are the best techniques for displaying the variables and their relationships?



# Audience and Purpose

- Audience:
  - Identify stakeholders
  - Understand user characteristics
  - Define user personas
  - Obtain user feedback



# Audience and Purpose

- Purpose
  - Identify business objectives or potential impact
  - Determine decision-making needs
  - Establish key questions
  - Prioritize information
  - Align with organizational goal or desired impact

# Key Message and Depth

- Clarify the primary (core) message
- Consider the audience
- Craft a central narrative
- Define the emotional impact
- Choose the right depth
- Work in layers
- Develop characters
- Employ symbols and metaphors
- Gather feedback and iterate

# Perform EDA

- Rely upon the dashboard objectives
- Understand the data sources and data available
- Assess the data quality
- Collect data
- Data integration, cleaning and transformation
- Descriptive Analysis
- Visualization

# Narrative

- Craft a persuasive and impactful presentation or report.
- Construct a narrative structure that guides the audience through the insights and information.
- This structure typically comprises:
  - An introduction to set the stage
  - A well-defined storyline to present the data coherently
  - A conclusion that ties together the key insights.
- The narrative should follow a logical flow, capturing the audience's attention and maintaining its engagement.

# Summary of the four Storytelling Principles

| Audience and Purpose | Exploratory Analysis |
|---|---|
| • Specific group or individuals: demographics, interests, and knowledge level.<br>• Clarify the objectives through the narrative. | • What questions do you want to answer with the data?<br>• What kind of relationships exist in the data?<br>• What are the best techniques for displaying the variables and their relationships? |
| **Key Message and Depth** | **Compelling Narrative** |
| • Central theme or insight: clear, concise, and relevant.<br>• Depth: level of detail, context, and supporting data. | • Persuasive and impactful presentation.<br>• Guide the audience through the insights and information. |

# Data Storytelling with AI

*“I have a dataset with the following variables: Row ID, Order ID, Order Date, Ship Date, Ship Mode, Customer ID, Customer Name, Segment, Country, City, State, Postal Code, Region, Product ID, Category, Sub-Category, Product Name, Sales. What types of analyses can I perform with such data and what types of insights I could get from the data?”*

# Data Storytelling with AI

*“Considering that my data storytelling workflow is composed of the steps below, please create a story around the dataset with the variables listed above.*

*1) Define audience and purpose*

*2) Identify key message and depth*

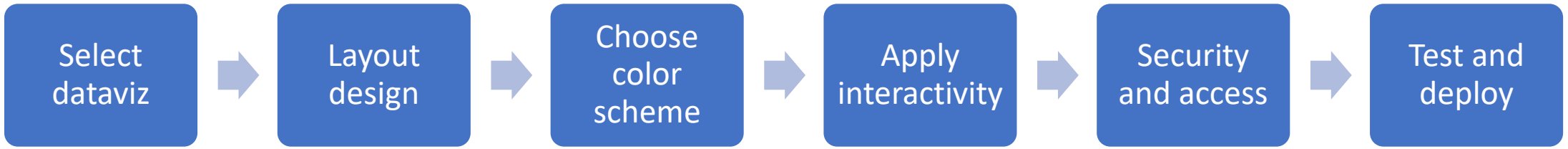
*3) Perform EDA*

*4) Prepare a compelling narrative”*

- Now input the dataset and prompt the model to create the story for this dataset.

# Dashboard Design

# Dashboard Design Process Flow



# Select Dataviz

- Consider the nature of data, because different types of data (categorical, numerical, time-series, text, graphs) require different visualizations.
- Take into account your audience.
- Keep visualizations simple and easy to interpret, avoiding clutter and choosing a format that minimizes confusion.
- Consider the three exploratory analysis guiding questions: What questions do you want to answer with the data? What are the relationships to be observed? What are the best methods to display the variables and their relationships?

# Design the Dashboard Layout

- The layout should follow a logical flow of information. This could be from left to right, top to bottom, or in any other pattern that suits the data and the story you are trying to tell.
- The arrangement should guide the user's eye through the dashboard in a meaningful sequence.
- The placement of charts and widgets on the dashboard should be carefully considered.
- Important information should be placed in prominent positions.
- Similar or related data should be grouped together.

# Design the Dashboard Layout

- The size of each chart or widget should reflect its importance, the complexity of the data it represents, or the general aesthetics of the dashboards.
- Strive for simplicity and clarity.
- Use space effectively to separate different sections or elements and avoid unnecessary decorations or elements that do not contribute to the understanding of the data, insights, and the story.

# Choose a Color Scheme

It is important to account for:

- Contrast and readability
- Color harmony
- Consistency
- Accessibility
- Limited palette
- Meaningful Color Choices

# Apply Interactivity

- **Filters**: Tools that allow users to limit the data that is displayed in a view. Types: categorical (e.g., filtering by product type or region), numerical (e.g., filtering by sales amount or quantity), or temporal (e.g., filtering by date or time). Usually presented as dropdown menus, checkboxes, or sliders.
- **Drill-Through**: Allows users to click on a data point in a chart or table and “drill through” to a more detailed report based on that data point. For example, you might click on a data point representing sales for a particular product and be taken to a different report showing detailed sales data for that product. The key aspect of drill-through is that it takes you to a different report or dashboard for more specific information.

# Apply Interactivity

- **Drill-Down**: Allows users to start with a high-level overview of data and then explore the data at a more granular level. For example, you might start by looking at sales data for an entire country, then drill down to see sales data for individual states, then counties, cities, and so on. The key aspect of drill-down is that you stay within the same report or dashboard, but the level of detail changes.
- **Tooltips**: Small information boxes that appear when you hover over an item on a dashboard, providing additional details about the item, such as exact values, labels, or explanations. This feature is particularly useful for presenting concise details without cluttering the main visualization.

# Sample Dashboards: Now, the end!

Real live dashboards (finance)

https://finance.yahoo.com/lookup

https://www.fundsexplorer.com.br/

https://www.marketwatch.com/

Companies that show live dashboards:

https://visual.is/examples

https://www.visualizefree.com/

https://www.klipfolio.com/live-dashboards

# Dashboard Design with AI

“*For the attached Superstore Sales Dataset recreate the story and design an interactive dashboard to extract insights from the data.*”

** “*I want you to create it in Power BI and a Plotly Dash app*.”

# Case Study

# Case Study: Superstore Sales

| Variable Name | Definition (meaning) | Variable Type | Range or Domain |
|---|---|---|---|
| Row ID | ID number of each row | Numerical (Integer) | Integer value (code) |
| Order ID | ID number of each order | String (alphanumeric) | {CA, US}-Year-integer value (code) |
| Order Date | Date product was ordered by the customer | Numerical (Integer) | Date |
| Ship Date | Date product was shipped to the customer | Numerical (Integer) | Date |
| Ship Mode | Class chosen for shipping | Categorical | {Same day, First Class, Second Class, Standard Class} |
| Customer ID | ID number for each customer | String (alphanumeric) | First and last initials of customer name}-integer value (code) |
| Customer Name | First and last name of customers | Categorical (Nominal) | {First and last name} |
| Segment | Category of customers | Categorical (Nominal) | {Consumer, Home office, Corporate} |
| Country | Customers' country of residence | Categorical (Nominal) | {United States} |
| City | Customers' city of residence | Categorical (Nominal) | City in the United States |
| State | Customers' state of residence | Categorical (Nominal) | State in the United States |
| Postal Code | Customers' postal code | Categorical (Integer) | Postal code in the United States |
| Region | Customers' region of residence in the US | Categorical (Nominal) | {Central, South, West, East} |
| Product ID | ID number of each product | String (alphanumeric) | Category-Sub-category-integer code |
| Category | Products' category | Categorical (Nominal) | {Office Supplies, Furniture, Technology} |
| Sub-Category | Type of products in each category | Categorical (Nominal) | {Accessories, Appliances, Art, Binders, Bookcases, Chairs, Copiers, Envelopes, Fasteners, Furnishings, Labels, Machines, Paper, Phones, Storage, Supplies, Tables} |
| Product Name | Names of products in each category | Categorical (Nominal) | Any product names |
| Sales | Total US$ by order | Numerical (Continuous) | [0.444, 22638.48] |

# Data Storytelling

1. **Audience and Purpose**: Individuals or groups involved in the decision-making process of a retail business, such as business analysts, sales managers, or executives.
2. **Key Message and Depth**: Understanding sales performance and trends to drive strategic decision-making in retail business. Focuses:
   - Sales Trends Over Time.
   - Sales by Category.
   - Geographical Distribution of Orders.
   - Overall Sales Performance.
   - Customizable Views.

## 3. Exploratory Analysis

### What questions do you want to answer with the data?

- How have sales values and the number of sales changed over time?
- Which categories and sub-categories have the highest sales values?
- How are orders distributed across different states?
- What is the total value of sales?
- How do sales values and the number of sales vary by year, shipment mode, and segment?

### What kinds of relationships exist in the data?

- Time (month/year) and total value or number of sales, indicating seasonal trends.
- Category/sub-category of products and the total sales values, suggesting some products are more popular or profitable than others.
- The number of orders might vary by state, indicating geographical trends in sales.
- The total value of sales could be influenced by factors such as the year, shipment mode, and segment.

## 4. Compelling Narrative

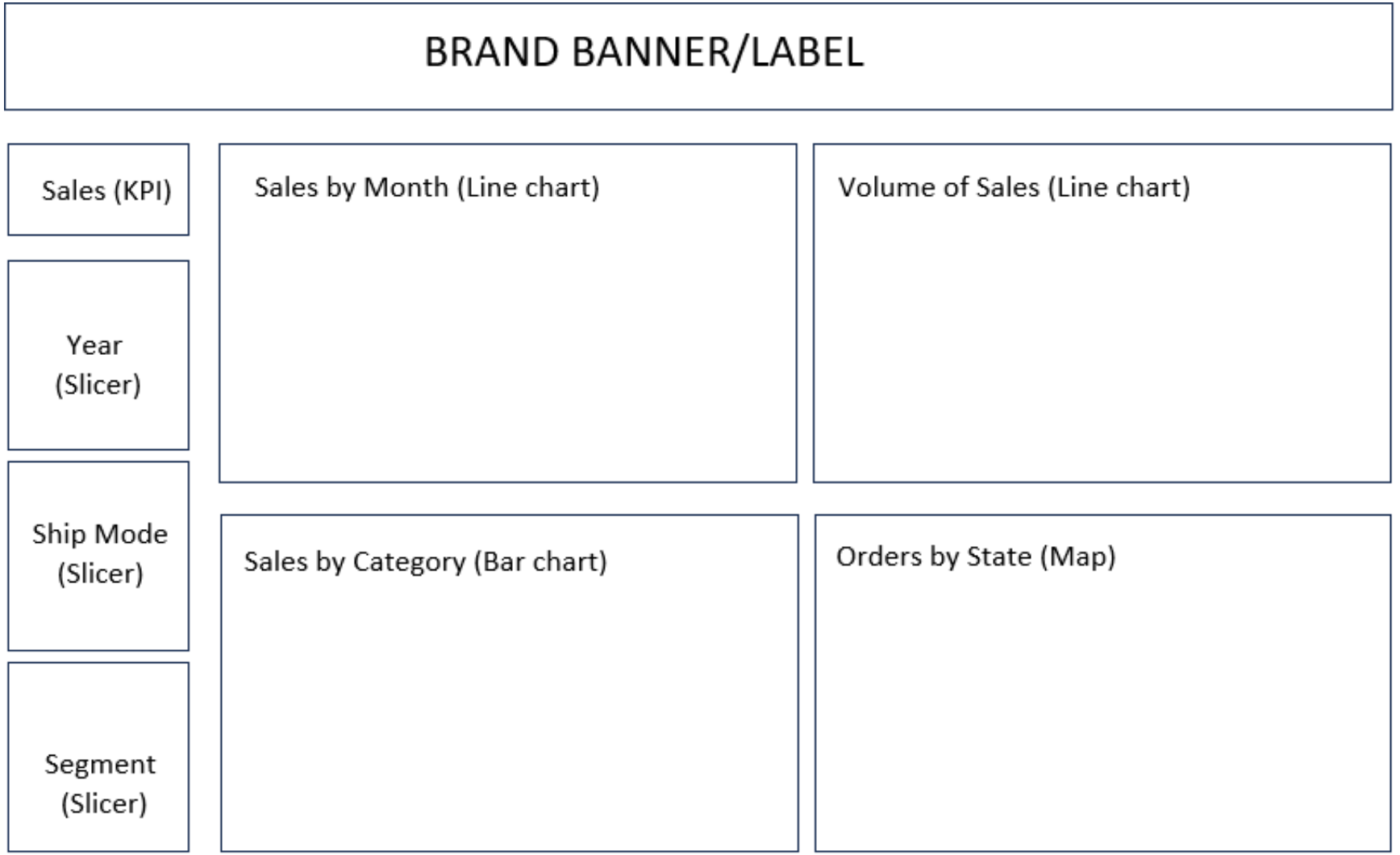
# Dashboard Design

## 1. Selecting Appropriate Visualizations

**What are the best techniques for displaying the variables and their relationships?**

- Line charts for displaying trends in total sales value and the number of sales over time.
- A bar chart for showing the total sales values by category and sub-category.
- A map for visualizing the number of orders by state.
- A KPI for presenting the total value of sales.
- Slicers for filtering the data by year, shipment mode, and segment, allowing the audience to explore the data in more depth.

# 2. Dashboard Layout



# 3. Choosing a Color Scheme

Different shades of green for the dashboard because it aligns well with several key principles of color selection for data visualization:

- Green is a color associated with positivity, growth, and harmony, making it suitable for conveying positive aspects of the data, such as sales growth or profitability.
- Green is recognized as a color representing “go” or “allowed”, making it intuitive for indicating positive values or actions, which is often the case in sales dashboards.
- Green provides excellent contrast against a white or light background, enhancing readability and ensuring that data points stand out effectively.

# 4. Applying Interactivity

- Slicers allow users to filter data based on specific time periods, shipment mode, and customer segments.
- The dashboard includes tooltips in all four visuals, which appear when hovering over data points, providing users with additional details such as exact values or labels without cluttering the main visualization.

# Superstore Sales Dashboard

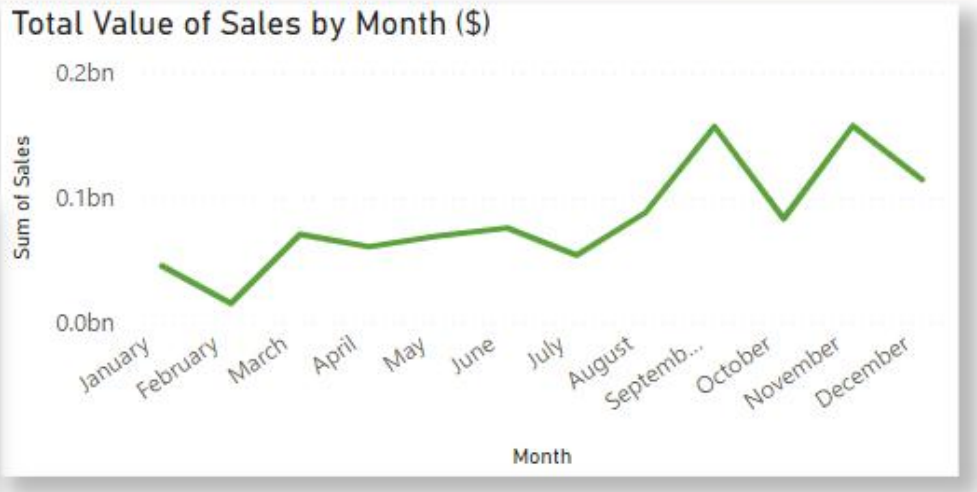
**Total Sales ($)**

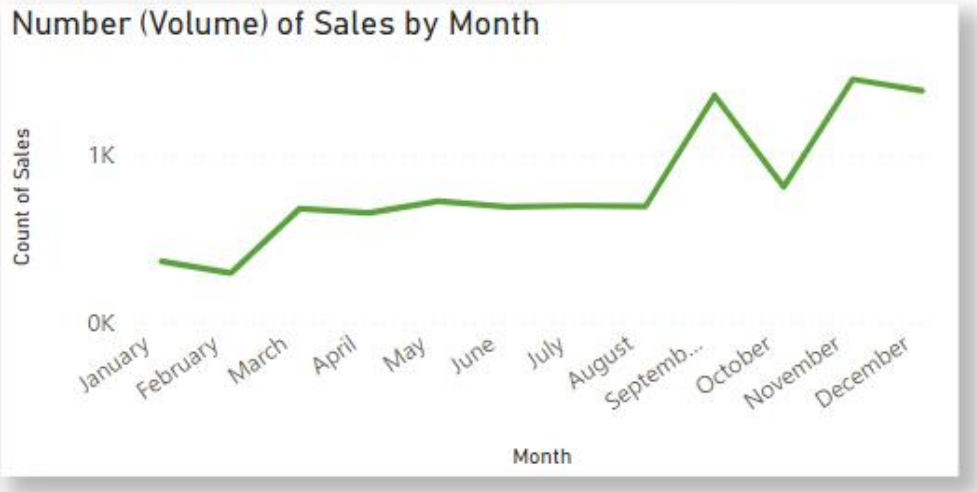
992.64M

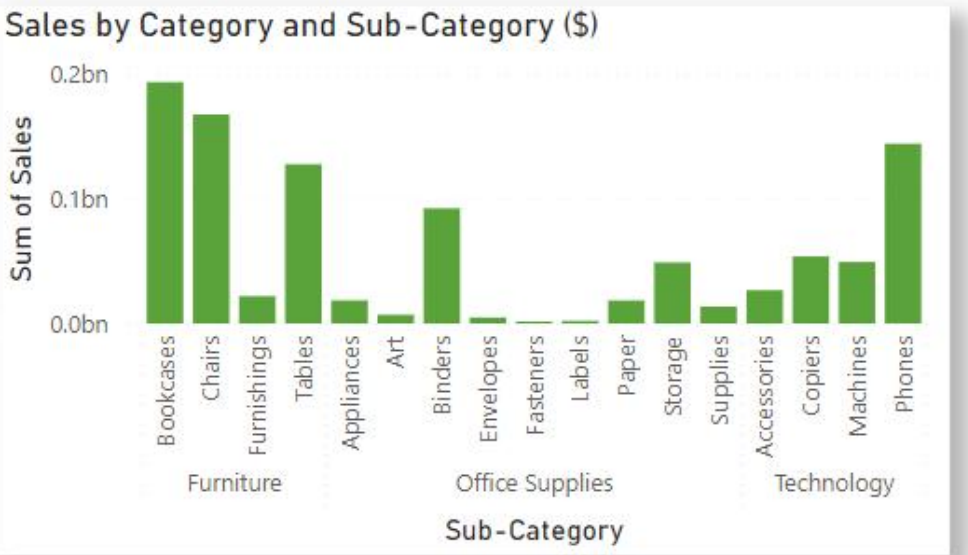
**Year**
- 2015
- 2016
- 2017
- 2018

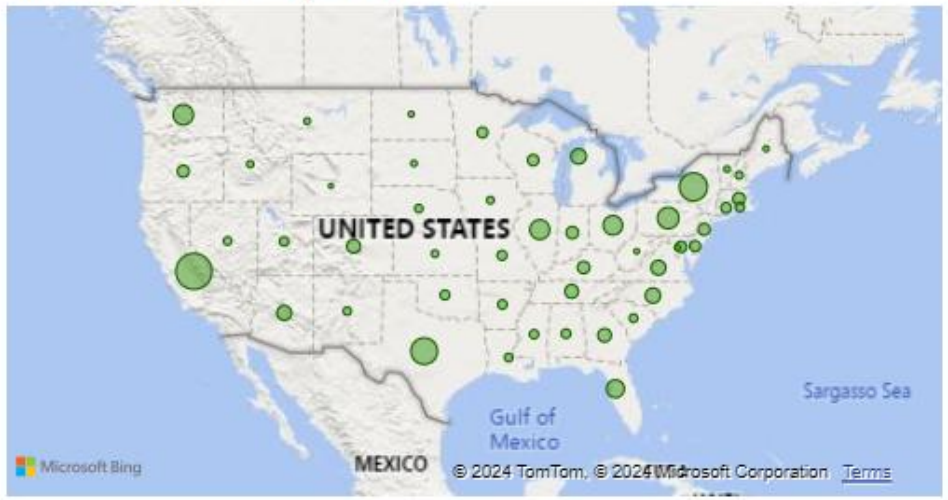
**Ship Mode**
- First Class
- Same Day
- Second Class
- Standard Class

**Segment**
- Consumer
- Corporate
- Home Office



# Leandro Nunes de Castro

**ldecastrosilva@fgcu.edu**