BSBINS401 - Analyse and Present Research Information

Session 2: Data Visualization Basics and Tools Overview

Lecturer: Jordan Hill

Learning Objectives

- Understand different types of graphs and when to use them
- Familiarize with key data visualization tools and platforms
- Learn basic functionalities of Jupyter Notebooks and Kaggle
- Get an overview of Matplotlib and Seaborn libraries

What is Data Visualization?

- **Definition**: The practice of translating information into a visual context
- **Purpose**:
  - Simplify the interpretation of complex data
  - Reveal patterns, trends, and correlations
  - Enhance data-driven decision making

Key Graph Types and Their Uses

Graph Type	Use Case
Line Chart	Trend over time
Bar Chart	Comparing discrete categories
Scatter Plot	Relationship between two variables
Histogram	Frequency distribution of data

Tools for Data Visualization

- **Jupyter Notebooks**
  - Interactive computing environment supporting live code, equations, visualizations, and narrative text
  - Jupyter Notebook Introduction Tutorial
- **Kaggle**
  - Platform for data science competitions and public datasets
  - Includes an in-browser coding environment
- **Visualization Libraries**
  - **Matplotlib**: Standard plotting library for Python (Matplotlib Getting Started)
  - **Seaborn**: High-level interface built on Matplotlib for statistical graphics (Seaborn Introduction)

Jupyter Notebook & Kaggle Overview

- **Jupyter Notebooks**:
  - Ideal for iterative development and data exploration
  - Combines code, output, and rich text in a single document
- **Kaggle**:
  - Create and share notebooks online
  - Explore datasets and participate in competitions
- **Demo**:
  - In-class demonstration of starting a notebook and navigating Kaggle

Introduction to Matplotlib and Seaborn

- **Matplotlib**:
  - Widely used for creating static plots
  - Offers detailed control over every aspect of a figure
- **Seaborn**:
  - Simplifies creating attractive statistical graphics
  - Provides high-level interfaces for drawing appealing graphs

Example Activity:

- Plot simple line charts and bar charts using sample code snippets

```
import matplotlib.pyplot as plt
import seaborn as sns

# Sample data
data = [1, 3, 2, 5, 7, 4]

# Matplotlib example
plt.plot(data)
plt.title("Simple Line Chart")
plt.show()

# Seaborn example
sns.barplot(x=list(range(len(data))), y=data)
plt.title("Simple Bar Chart")
plt.show()
```

In-Class Activities

- **Explore Jupyter Notebooks and Kaggle**:
  - Create a new notebook and run a "Hello, World!" cell
  - Browse Kaggle to locate datasets and review notebook examples
- **Experiment**:
  - Activity: Explore Different Types of Graphs
  - Modify code snippets (colors, labels, etc.) to create your own visualizations

Activities

- Activity: Exploring Different Types of Graphs
- Jupyter Notebook: Introduction Tutorial

Lab Resources

- Lab: Markdown in Jupyter Notebooks
- Lab Exercise: Simple Plotting with Matplotlib

Additional Resources

- **Matplotlib Documentation**: Getting Started
- **Seaborn Documentation**: Introduction
- **Kaggle Learn**: Intro to Data Visualization
- **Jupyter Notebook Tutorials**: Jupyter Docs

Further Reading

- **Article:** Data Visualization Best Practices

Summary & Next Steps

- **Today's Session**:
  - Reviewed key graph types, visualization tools, and platforms
  - Demonstrated usage of Jupyter Notebook and Kaggle
- **Next Session Preview**:
  - Dive deeper into data visualization techniques
  - Engage in more hands-on plotting exercises and explore advanced customization options

Any Questions?

- Please ask any questions now or feel free to reach out during office hours.
- Contact: jordan.hill@nmtafe.wa.edu.au