# Communicating Data Insights

## COMMUNICATING INSIGHTS FROM STATISTICAL DATA

**Janani Ravi**
CO-FOUNDER, LOONYCORN

www.loonycorn.com

# Overview

Working with Matplotlib visualizations

Box plots and violin plots

Viewing frequency distributions using histograms

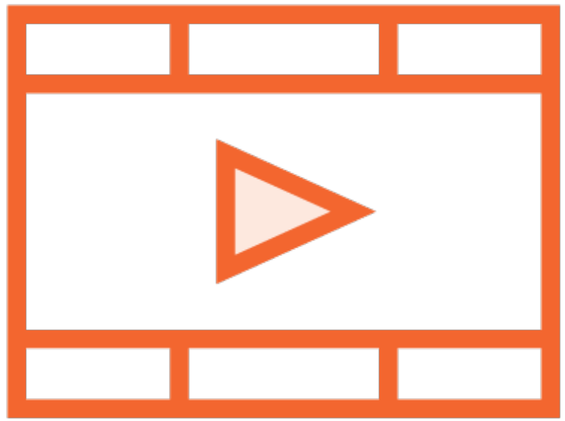
Viewing composition using pie charts

Autocorrelation in time series data

Stacked plots and stem plots

# Prerequisites and Course Outline

# Prerequisites



**Basic Python programming**

**Basic SQL queries**

**Some familiarity work on cloud platforms i.e. Azure**

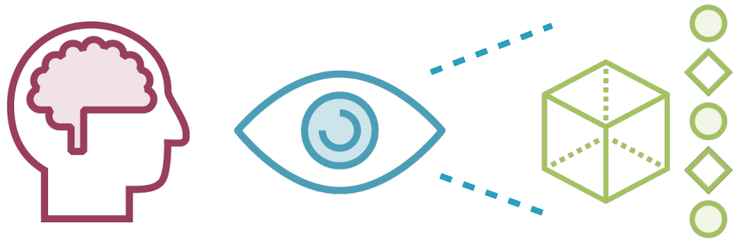
**High school math**

# Course Outline

- Drawing insights from statistical data
- Drawing insights from business data
- Visualizing relationships and distributions
- Integrating data in multi-cloud environments
- Integrating data in hybrid environments

# Visualization Libraries in Python

# Visualization in Exploratory Data Analysis



**Important step in data exploration**

**Helps developing an intuition for relationships in data**

**Precursor to higher level data analysis using ML techniques**

# Interactivity

**Interactivity helps with exploration and experimentation**

# Interactivity in Visualization

**Easy to underestimate importance of interactivity**

**Lacking from many visualization tools**

**Enables exploration**

**Dramatically increases understanding**

# Visualization Libraries in Python

**Matplotlib**

**Seaborn**

**Bokeh**

**Plotly.py**

# Many Libraries, Many Niches

**Matplotlib is powerful**

**Seaborn is easy-to-use**

**Bokeh for interactivity**

**Plotly.py for collaboration**

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| Pie charts | Parts of a whole |
| Line charts | Changes over time |
| Bar graphs | Same quantity for different groups |
| Histograms | Distribution of data across bins |
| Stacked bar charts | Parts of whole as well as comparison across groups |

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| Scatter plots | Co-movement, outlier detection |
| Box plots | Quartiles (median, range, IQR) |
| Violin plots | Similar to box plots but also show probability density |

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| **Rug Plots** | **One-dimensional scatter plot, also zero bin-width histogram** |

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| KDE plots | Smoothened probability distribution |
| Violin plots | Combine Box plot and KDE plot |

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| **Sankey diagrams** | **Flow information e.g. Napoleon's invasion of Russia** |

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| **Funnel plots** | **Sequential step-by-step processes with loss at each step** |

# Picking the Right Visualization Type

| Visualization Type | Use Case |
| --- | --- |
| Candlestick plots | Stock price movements |

# Demo

## Visualizing statistical data

# Demo

**Box plots and violin plots**

# Demo

## Histograms

# Demo

**Pie charts**

# Demo

## Visualizing autocorrelation

# Demo

**Stacked plots and stem plots**

# Summary

- Working with Matplotlib visualizations
- Box plots and violin plots
- Viewing frequency distributions using histograms
- Viewing composition using pie charts
- Autocorrelation in time series data
- Stacked plots and stem plots