Rainfall in Pakistan (1901–2016)

Exploratory Data Analysis

SMIT - AI and Data Science

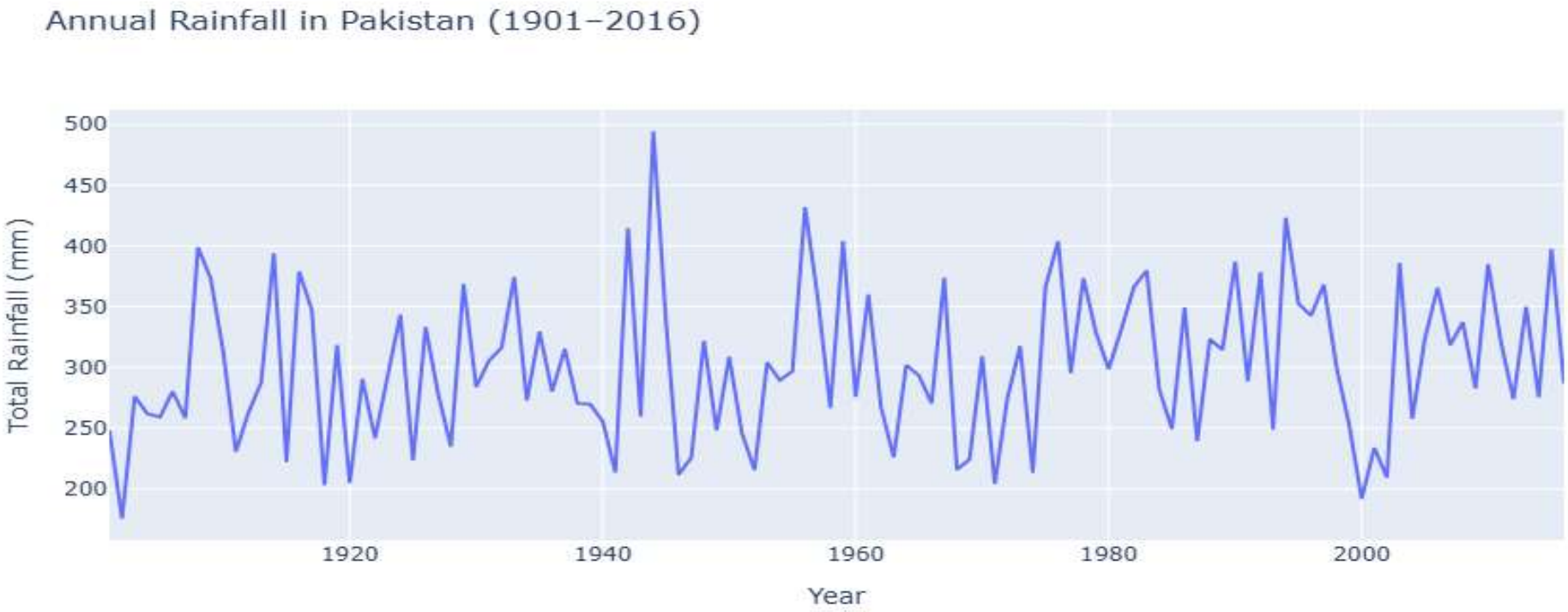
Introduction:

- **Dataset:** Rainfall 1901–2016
- **Dataset Link:** https://www.kaggle.com/datasets/zusmani/rainfall-in-pakistan
- **Objective:** Explore patterns, trends, and climate change impact

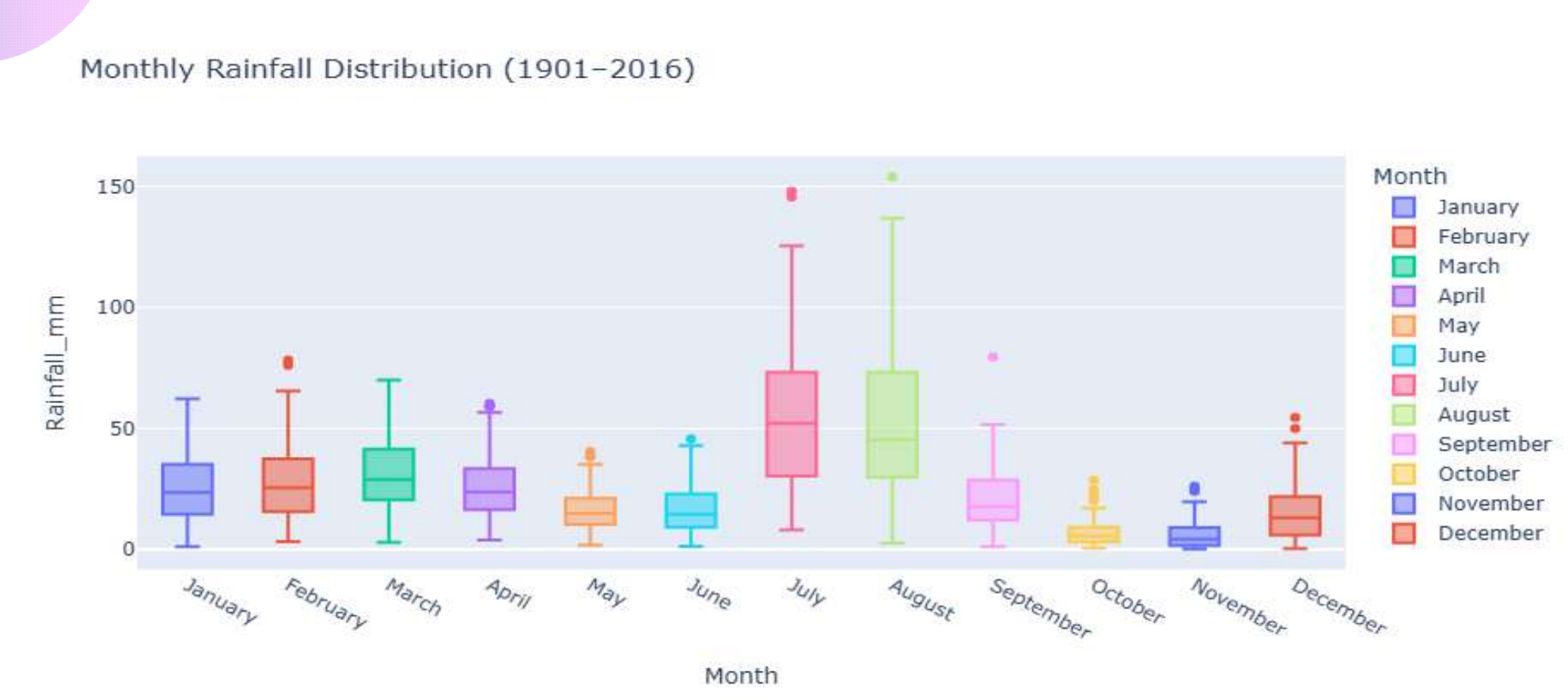
Dataset Overview:

- **Columns:** Year, Month, Rainfall_mm
- **Shape:** (1392, 3)
- **Missing Values: Rainfall_mm** = 0, **Year** = 0, **Month** = 0

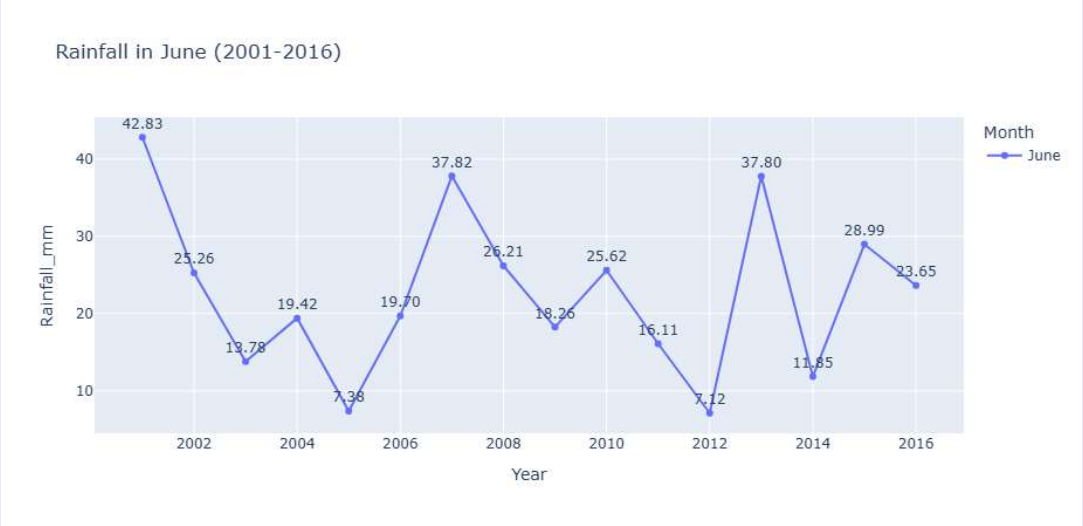
Annual Rainfall Trends



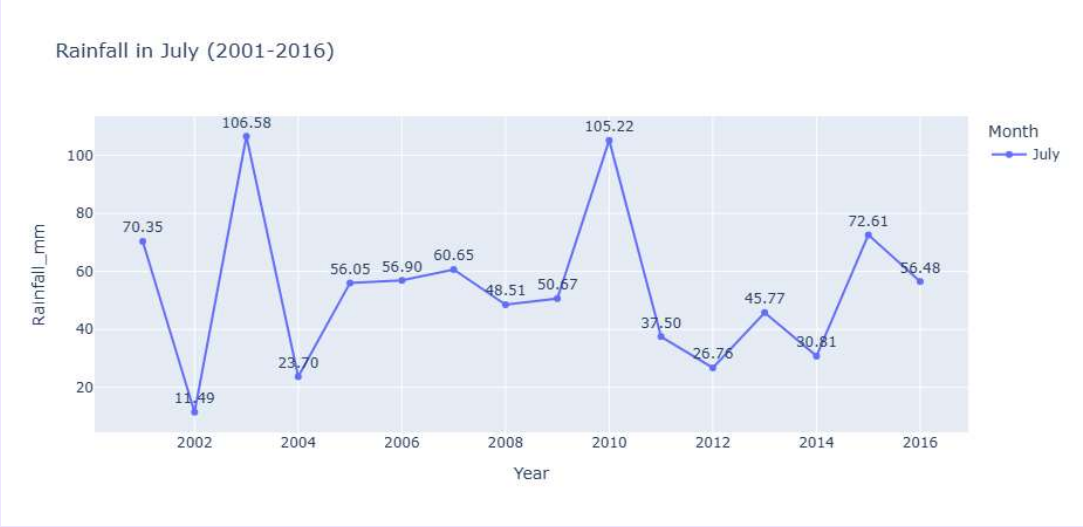
Monthly Rainfall Distribution



Monsoon Trends (June)

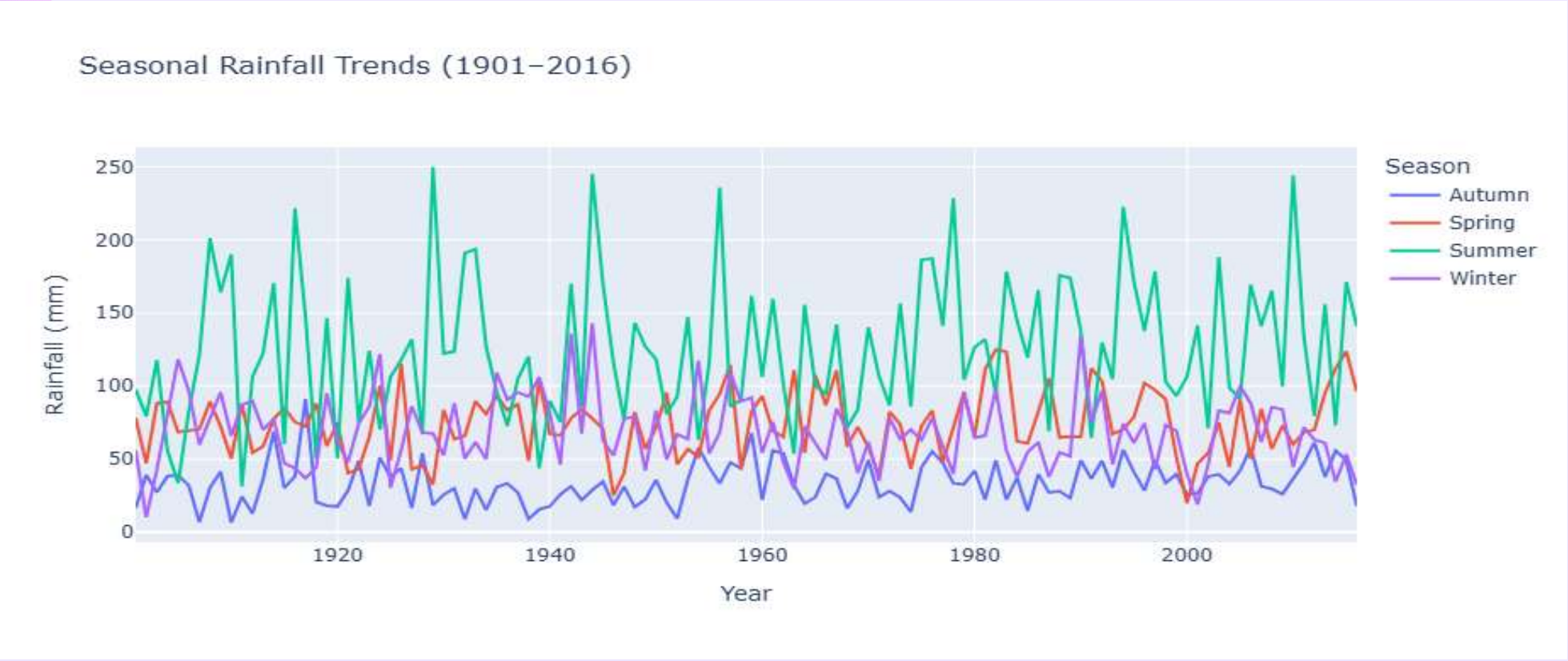


Monsoon Trends (July)

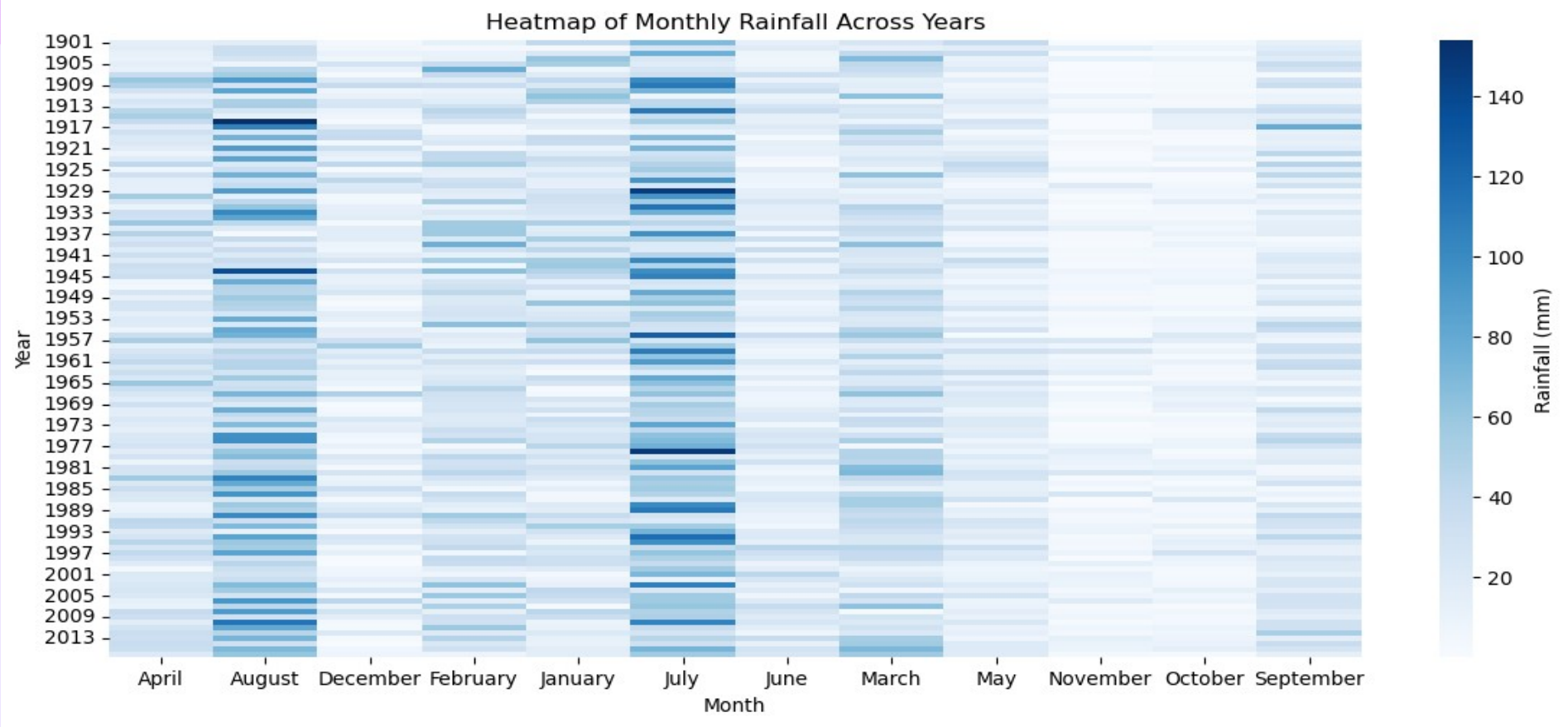


Monsoon Trends (August)

Seasonal Rainfall Trends



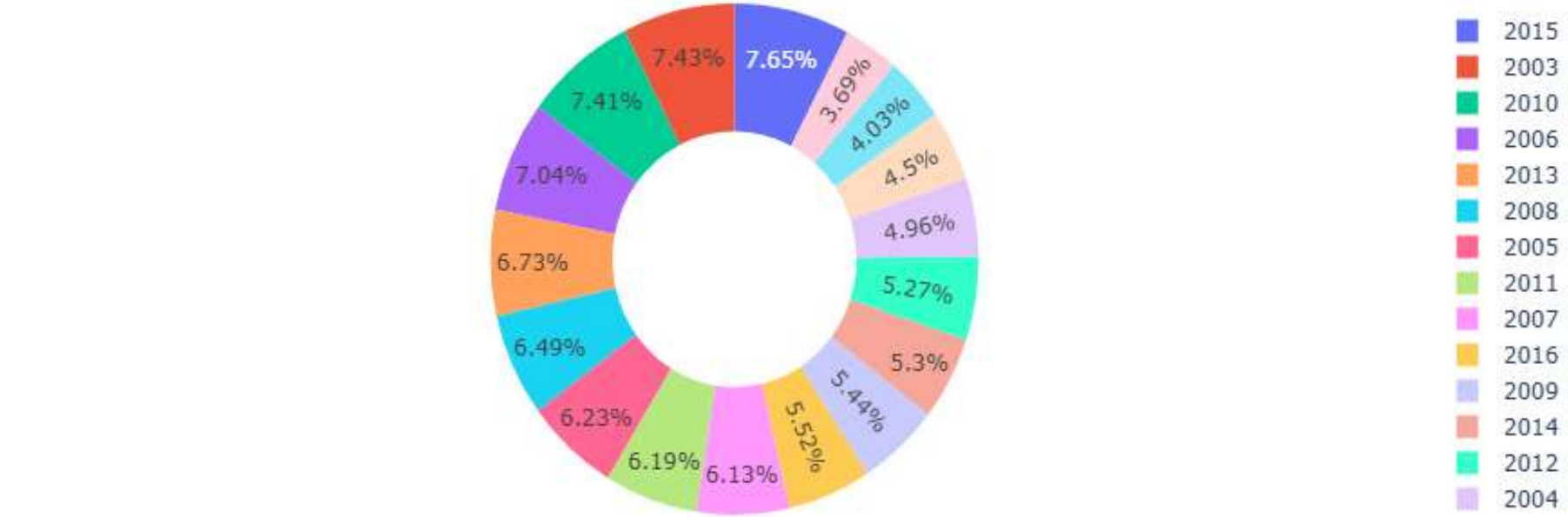
Heatmap of Rainfall



Hypothesis Testing

```
Hypothesis Testing (Before 2000 vs After 2000)
T-statistic: -0.28845512310745 85
P-value: 0.7735222001256075
No significant difference in rainfall before and after 2000.
```

Rainfall Contribution (2000–2016)



Conclusion

Rainfall trends in Pakistan show long-term variability, with most rainfall occurring during the monsoon months (June–August). Hypothesis testing shows no significant difference in rainfall before and after 2000. Summer receives the highest rainfall, driven by monsoon activity. These findings highlight the need for continued climate monitoring and improved water resource management.