# Rainfall Analysis in Pakistan (1901–2016)

Author: Faraz Abbasi
Date: September 14, 2025

## 1. Introduction

This project performs an Exploratory Data Analysis (EDA) of rainfall data in Pakistan from 1901 to 2016. The analysis aims to identify trends, patterns, and anomalies in rainfall across years, months, and seasons. It also evaluates whether there has been a statistically significant change in rainfall patterns after the year 2000.

## 2. Dataset Overview

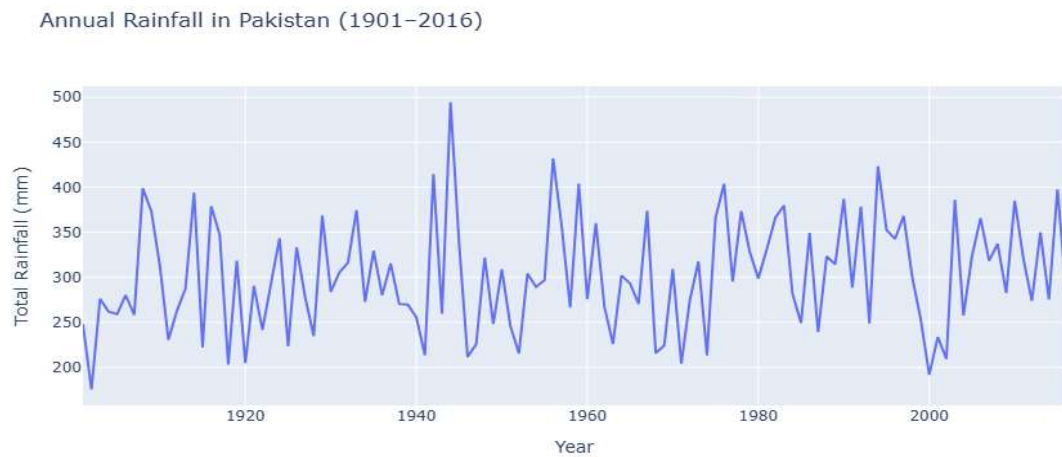**Source:** Rainfall dataset (1901–2016) for Pakistan

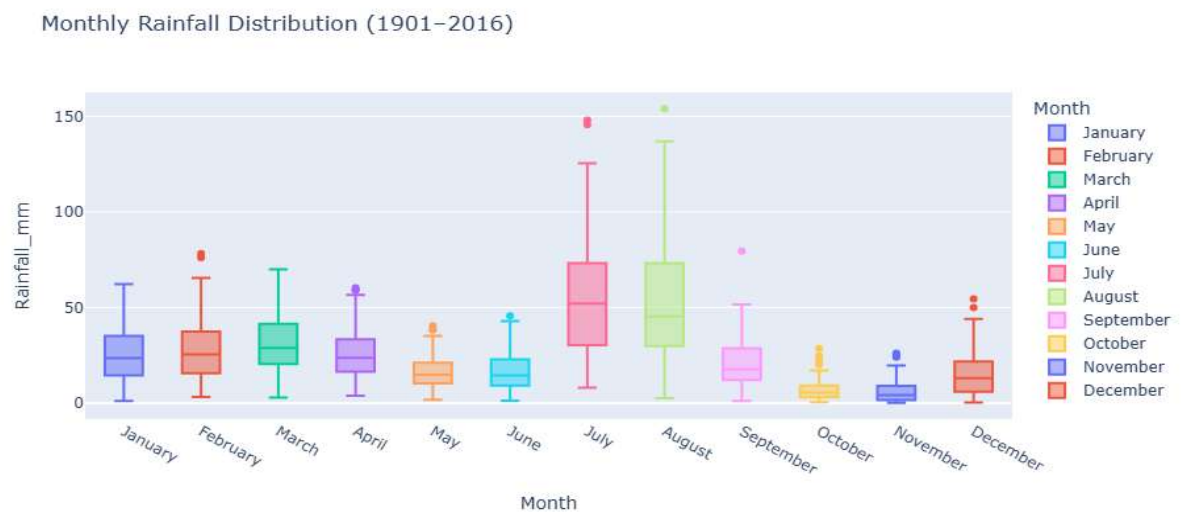**Columns:** Year, Month, Rainfall_mm

**Shape:** (1392, 3)

**Missing Values:** Rainfall_mm = 0, Year = 0, Month = 0

## 3. Annual Rainfall Trends

We calculated total annual rainfall for each year.

Annual Rainfall in Pakistan (1901–2016)

## 4. Monthly Rainfall Distribution

Monthly Rainfall Distribution (1901–2016)
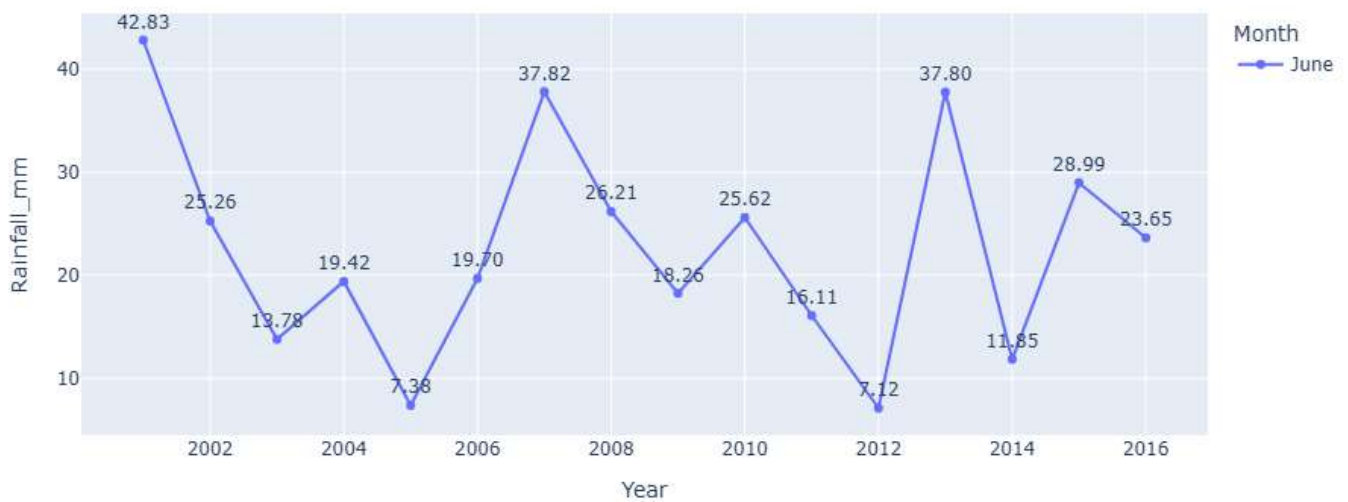
## 5. Monthly Trends (Focus on Monsoon Months)

We analyzed rainfall trends for June, July, and August before and after 2000.
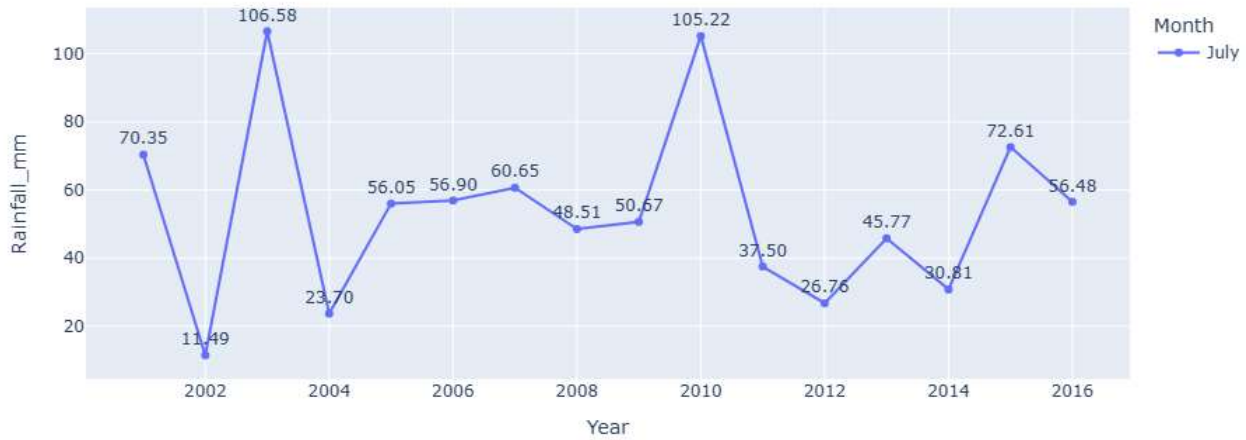
**June:**

### Rainfall in June (1980-2000)



### Rainfall in June (2001-2016)

**July:**

### Rainfall in July (2001-2016)
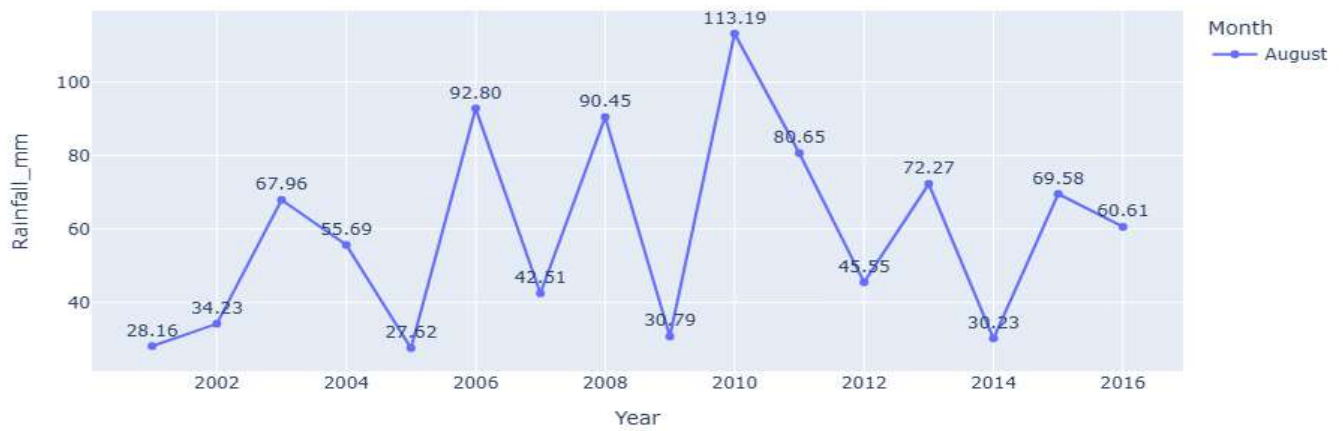


### Rainfall in July (1980-2000)

**August:**

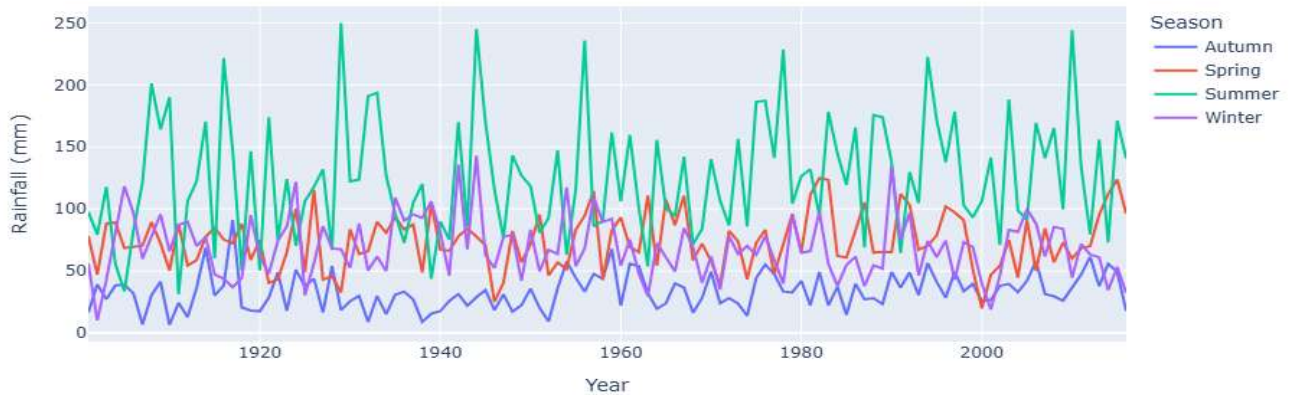### Rainfall in August (1980-2000)



### Rainfall in August (2001-2016)
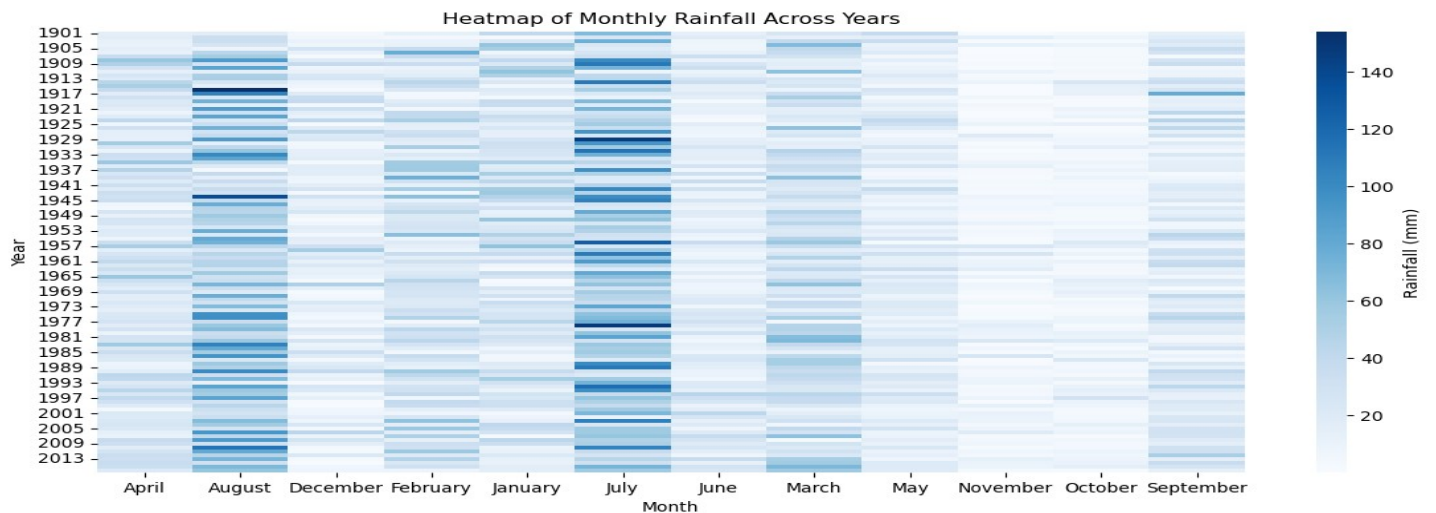
## 6. Seasonal Rainfall

Rainfall was grouped into Winter, Summer, Spring, and Autumn.



Seasonal Rainfall Trends (1901–2016)

## 7. Heatmap of Monthly Rainfall

We generated a heatmap to visualize monthly rainfall variation across years.



Heatmap of Monthly Rainfall Across Years

## 8. Hypothesis Testing

● Null Hypothesis ($H_0$): No difference in rainfall before and after 2000.

● Alternative Hypothesis ($H_1$): There is a difference.

```
Hypothesis Testing (Before 2000 vs After 2000)
T-statistic: -0.2884551231074585
P-value: 0.773522201256075
No significant difference in rainfall before and after 2000.
```
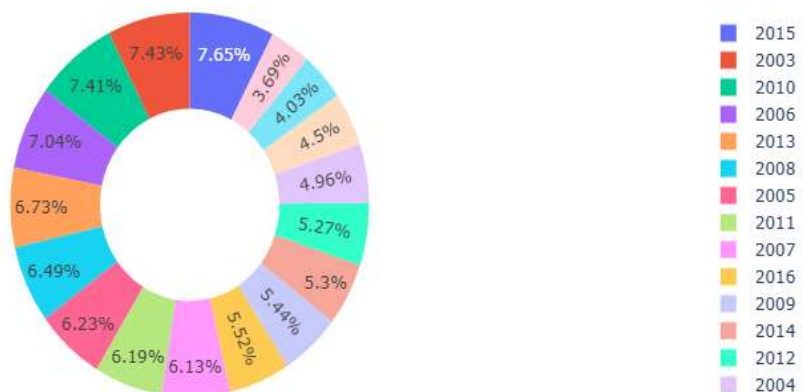
### Interpretation:
If p-value < 0.05: Significant difference in rainfall.

If p-value ≥ 0.05: No significant difference.

## 9. Rainfall Contribution (2000–2016)
Pie chart showing yearly contributions after 2000.



Rainfall Contribution by Year (2000–2016)

## 10. 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.