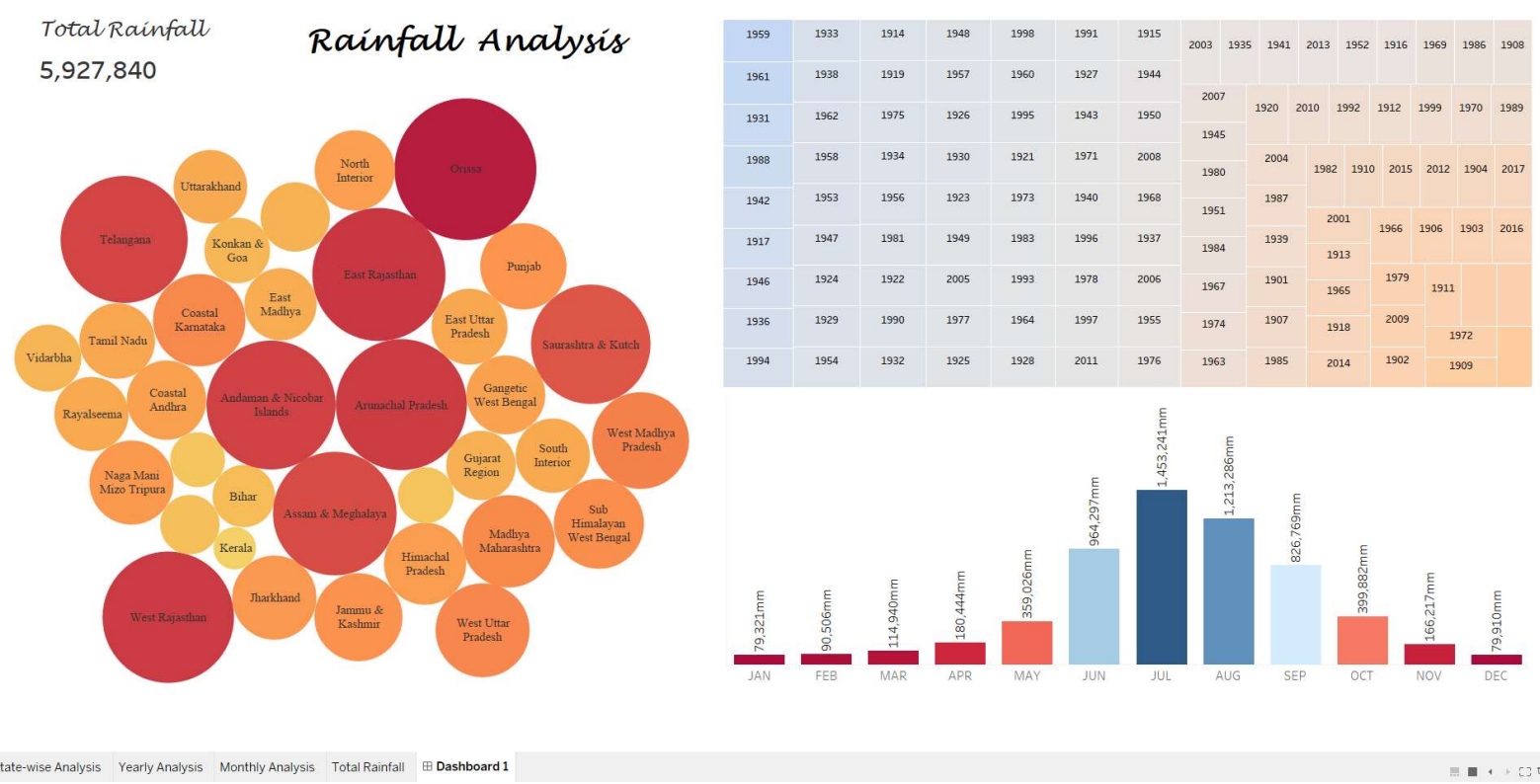# Foundations of Data Science (UCS548)
LAB2 Evaluation
TABLEAU DASHBOARD

Name: Tanishq Singla
Roll No.: 102003205
Group: 3COE9

- Dashboard

- ## Datasets
- ## Dataset 1 (Annual Data)

	States/UTs	YEAR	ANNUAL
1	States/UTs	YEAR	ANNUAL
2	Andaman & Nicobar Islands	1901	3373.2
3	Andaman & Nicobar Islands	1902	3520.7
4	Andaman & Nicobar Islands	1903	2957.4
5	Andaman & Nicobar Islands	1904	3079.6
6	Andaman & Nicobar Islands	1905	2566.7
7	Andaman & Nicobar Islands	1906	2534.4
8	Andaman & Nicobar Islands	1907	3347.9
9	Andaman & Nicobar Islands	1908	3576.4
10	Andaman & Nicobar Islands	1910	2899.4
11	Andaman & Nicobar Islands	1911	2687.2
12	Andaman & Nicobar Islands	1912	2960.5
13	Andaman & Nicobar Islands	1913	2365.8
14	Andaman & Nicobar Islands	1914	2957.8
15	Andaman & Nicobar Islands	1915	2741.3
16	Andaman & Nicobar Islands	1916	2937.5
17	Andaman & Nicobar Islands	1917	2612.4
18	Andaman & Nicobar Islands	1918	3275
19	Andaman & Nicobar Islands	1919	2352.1
20	Andaman & Nicobar Islands	1920	2943.2
21	Andaman & Nicobar Islands	1921	2606.4
22	Andaman & Nicobar Islands	1922	3554.2
23	Andaman & Nicobar Islands	1923	NA
24	Andaman & Nicobar Islands	1924	NA
25	Andaman & Nicobar Islands	1925	2480.5
26	Andaman & Nicobar Islands	1926	3282.2
27	Andaman & Nicobar Islands	1927	2442.9
28	Andaman & Nicobar Islands	1928	2998.3
29	Andaman & Nicobar Islands	1929	2926.6

- ## Dataset 2 (First 6 Months Data)

	States/UTs	YEAR	JAN	FEB	MAR	APR	MAY	JUN
1	States/UTs	YEAR	JAN	FEB	MAR	APR	MAY	JUN
2	Andaman & Nicobar Islands	1901	49.2	87.1	29.2	2.3	528.8	517.5
3	Andaman & Nicobar Islands	1902	0	159.8	12.2	0	446.1	537.1
4	Andaman & Nicobar Islands	1903	12.7	144	0	1	235.1	479.9
5	Andaman & Nicobar Islands	1904	9.4	14.7	0	202.4	304.5	495.1
6	Andaman & Nicobar Islands	1905	1.3	0	3.3	26.9	279.5	628.7
7	Andaman & Nicobar Islands	1906	36.6	0	0	0	556.1	733.3
8	Andaman & Nicobar Islands	1907	110.7	0	113.3	21.6	616.3	305.2
9	Andaman & Nicobar Islands	1908	20.9	85.1	0	29	562	693.6
10	Andaman & Nicobar Islands	1910	26.6	22.7	206.3	89.3	224.5	472.7
11	Andaman & Nicobar Islands	1911	0	8.4	0	122.5	327.3	649
12	Andaman & Nicobar Islands	1912	583.7	0.8	0	21.9	140.7	549.8
13	Andaman & Nicobar Islands	1913	84.8	0.5	1.3	2.5	190.7	530
14	Andaman & Nicobar Islands	1914	0	0	0	37.7	298.8	383.3
15	Andaman & Nicobar Islands	1915	45	56.7	33.3	40.9	170.2	334.7
16	Andaman & Nicobar Islands	1916	0	0	0	0.5	487.4	450.1
17	Andaman & Nicobar Islands	1917	8	3.6	112	4.5	295.9	301.1
18	Andaman & Nicobar Islands	1918	77.4	6.9	11.4	10.7	729.3	710.8
19	Andaman & Nicobar Islands	1919	10.2	18	0	35.5	283.9	542.5
20	Andaman & Nicobar Islands	1920	122.3	7.4	3.1	13	237.4	546.9
21	Andaman & Nicobar Islands	1921	13.2	3.1	0	37.5	351.2	282.7
22	Andaman & Nicobar Islands	1922	245.3	34.3	15.6	323.1	289.7	506.1
23	Andaman & Nicobar Islands	1923	79.5	0	NA	91.3	293.5	808.4
24	Andaman & Nicobar Islands	1924	28.7	0	14.8	89.7	191.2	261.2
25	Andaman & Nicobar Islands	1925	36.6	0	8.6	50.4	282.2	663.8
26	Andaman & Nicobar Islands	1926	122.1	0	0	0.5	198.4	370
27	Andaman & Nicobar Islands	1927	3	17.5	17.8	108.6	504.1	433.3
28	Andaman & Nicobar Islands	1928	50.9	67.6	80.7	129.3	499.5	410.2
29	Andaman & Nicobar Islands	1929	74.2	118.4	129.2	69.8	316.6	588.8

- Dataset 3 (Last 6 Months Data)

States/UTs	YEAR	JUL	AUG	SEP	OCT	NOV	DEC
Andaman & Nicobar Islands	1901	365.1	481.1	332.6	388.5	558.2	33.6
Andaman & Nicobar Islands	1902	228.9	753.7	666.2	197.2	359	160.5
Andaman & Nicobar Islands	1903	728.4	326.7	339	181.2	284.4	225
Andaman & Nicobar Islands	1904	502	160.1	820.4	222.2	308.7	40.1
Andaman & Nicobar Islands	1905	368.7	330.5	297	260.7	25.4	344.7
Andaman & Nicobar Islands	1906	247.7	320.5	164.3	267.8	128.9	79.2
Andaman & Nicobar Islands	1907	443.9	377.6	200.4	264.4	648.9	245.6
Andaman & Nicobar Islands	1908	481.4	699.9	428.8	170.7	208.1	196.9
Andaman & Nicobar Islands	1910	264.3	337.4	626.6	208.2	267.3	153.5
Andaman & Nicobar Islands	1911	253	187.1	464.5	333.8	94.5	247.1
Andaman & Nicobar Islands	1912	468.9	370.3	386.2	318.7	117.2	2.3
Andaman & Nicobar Islands	1913	280.8	205.8	580.1	288.8	133	67.5
Andaman & Nicobar Islands	1914	792.8	520.5	310.8	139.8	184.4	289.7
Andaman & Nicobar Islands	1915	269	317.2	429.8	468.1	258.4	318
Andaman & Nicobar Islands	1916	317.3	425	561.2	369.7	192.6	133.7
Andaman & Nicobar Islands	1917	394.8	437.4	471.8	238.1	108.3	236.9
Andaman & Nicobar Islands	1918	200.9	455.4	303.3	227	366.9	175
Andaman & Nicobar Islands	1919	246.5	259.8	170.7	186.2	340.4	258.4
Andaman & Nicobar Islands	1920	294.4	467.4	505.4	397.5	262.9	85.5
Andaman & Nicobar Islands	1921	487.1	330	581.2	360.7	118.2	41.5
Andaman & Nicobar Islands	1922	425.8	307.4	511.7	162	541	192.2
Andaman & Nicobar Islands	1923	636.9	182.2	560.5	131.9	197.4	70.6
Andaman & Nicobar Islands	1924	493.3	290.9	251.2	331.1	378.6	NA
Andaman & Nicobar Islands	1925	241.8	278.2	201.9	249.5	271.5	196
Andaman & Nicobar Islands	1926	195.3	523.7	719.3	443.8	148.4	560.7
Andaman & Nicobar Islands	1927	195.2	370.1	126.2	327.5	274.1	65.5
Andaman & Nicobar Islands	1928	406.3	391.5	404.8	444.5	99.5	13.5
Andaman & Nicobar Islands	1929	134	644.7	172.9	413	251.5	13.5

- Data Clearing and Refining Steps
1. Grouping dataset with respect to a column.
2. Summarising the dataset with respect to mean of a column.
3. Removing na values with mean of the corresponding row and column.
4. Generating new columns by using other column values.
5. Merging datasets to create the final dataset.

```r
# Importing required libraries
library(dplyr)
library(writexl)

# Loading the datasets
full <- read.csv('Dataset/Dataset1.csv')
half1 <- read.csv('Dataset/Dataset2.csv')
half2 <- read.csv('Dataset/Dataset3.csv')

# Grouping January to June data based on Stated.UTs
df_half1 <- group_by(half1, States.UTs)

# Replacing na values by mean
df_Jan <- summarise(df_half1, JAN = mean(JAN, na.rm = T))
for(i in df_Jan$States.UTs)
{
  a <- is.na(half1$JAN)
  b <- half1$States.UTs==i
  half1$JAN[a&b] <- round(df_Jan$JAN[df_Jan$States.UTs==i],2)
}

df_Feb <- summarise(df_half1, FEB = mean(FEB, na.rm = T))
for(i in df_Feb$States.UTs)
{
  a <- is.na(half1$FEB)
  b <- half1$States.UTs==i
  half1$FEB[a&b] <- round(df_Feb$FEB[df_Feb$States.UTs==i],2)
}

df_Mar <- summarise(df_half1, MAR = mean(MAR, na.rm = T))
for(i in df_Mar$States.UTs)
{
  a <- is.na(half1$MAR)
  b <- half1$States.UTs==i
  half1$MAR[a&b] <- round(df_Mar$MAR[df_Mar$States.UTs==i],2)
}

df_Apr <- summarise(df_half1, APR = mean(APR, na.rm = T))
for(i in df_Mar$States.UTs)
{
  a <- is.na(half1$APR)
  b <- half1$States.UTs==i
  half1$APR[a&b] <- round(df_Apr$APR[df_Apr$States.UTs==i],2)
}

df_May <- summarise(df_half1, MAY = mean(MAY, na.rm = T))
for(i in df_May$States.UTs)
{
  a <- is.na(half1$MAY)
  b <- half1$States.UTs==i
  half1$MAY[a&b] <- round(df_May$MAY[df_May$States.UTs==i],2)
}

df_Jun <- summarise(df_half1, JUN = mean(JUN, na.rm = T))
for(i in df_Jun$States.UTs)
{
  a <- is.na(half1$JUN)
  b <- half1$States.UTs==i
  half1$JUN[a&b] <- round(df_Jun$JUN[df_Jun$States.UTs==i],2)
}

# Grouping July to December data based on Stated.UTs
df_half2 <- group_by(half2, States.UTs)

# Replacing na values by mean
df_Jul <- summarise(df_half2, JUL = mean(JUL, na.rm = T))
for(i in df_Jul$States.UTs)
{
  a <- is.na(half2$JUL)
  b <- half2$States.UTs==i
  half2$JUL[a&b] <- round(df_Jul$JUL[df_Jul$States.UTs==i],2)
}

df_Aug <- summarise(df_half2, AUG = mean(AUG, na.rm = T))
for(i in df_Aug$States.UTs)
{
  a <- is.na(half2$AUG)
  b <- half2$States.UTs==i
  half2$AUG[a&b] <- round(df_Aug$AUG[df_Aug$States.UTs==i],2)
}
```

```
 82  df_Sep <- summarise(df_half2, SEP = mean(SEP, na.rm = T))
 83  for(i in df_Sep$States.UTs)
 84 ▾ {
 85    a <- is.na(half2$SEP)
 86    b <- half2$States.UTs==i
 87    half2$SEP[a&b] <- round(df_Sep$SEP[df_Sep$States.UTs==i],2)
 88 ▴ }
 89
 90  df_Oct <- summarise(df_half2, OCT = mean(OCT, na.rm = T))
 91  for(i in df_Oct$States.UTs)
 92 ▾ {
 93    a <- is.na(half2$OCT)
 94    b <- half2$States.UTs==i
 95    half2$OCT[a&b] <- round(df_Oct$OCT[df_Oct$States.UTs==i],2)
 96 ▴ }
 97
 98  df_Nov <- summarise(df_half2, NOV = mean(NOV, na.rm = T))
 99  for(i in df_Nov$States.UTs)
100 ▾ {
101    a <- is.na(half2$NOV)
102    b <- half2$States.UTs==i
103    half2$NOV[a&b] <- round(df_Nov$NOV[df_Nov$States.UTs==i],2)
104 ▴ }
105
106  df_Dec <- summarise(df_half2, DEC = mean(DEC, na.rm = T))
107  for(i in df_Dec$States.UTs)
108 ▾ {
109    a <- is.na(half2$DEC)
110    b <- half2$States.UTs==i
111    half2$DEC[a&b] <- round(df_Dec$DEC[df_Dec$States.UTs==i],2)
112 ▴ }
113
114  # Merging first six months with last six months
115  df1 <- merge(x=half1, y=half2, by = c("States.UTs", "YEAR"),all.x=TRUE)
116
117  # Removing na values in the full(yearly) dataset
118  # by replacing it with sum of months of df1
119  full$ANNUAL <- apply(df1[,3:14], 1, sum)
120
121  # Generating new columns (Quarters of the year)
122  df1$FirstQuarter <- apply(df1[,3:5], 1, sum)
123
124  df1$SecondQuarter <- apply(df1[,6:8], 1, sum)
125
126  df1$ThirdQuarter <- apply(df1[,9:11], 1, sum)
127
128  df1$FourthQuarter <- apply(df1[,12:14], 1, sum)
129
130  # Merging all the datasets
131  df <- merge(x=df1, y=full, by = c("States.UTs", "YEAR"),all.x=TRUE)
132
133  # Writing the final dataset to an excel file
134  write_xlsx(df,"FinalData.xlsx")
135
```

- Summarised Tables

```
> head(df_Jan)                          > head(df_Aug)
# A tibble: 6 × 2                        # A tibble: 6 × 2
  States.UTs                   JAN         States.UTs                   AUG
  <chr>                       <dbl>        <chr>                       <dbl>
1 Andaman & Nicobar Islands  54.4       1 Andaman & Nicobar Islands  399.
2 Arunachal Pradesh          46.7       2 Arunachal Pradesh          491.
3 Assam & Meghalaya          16.9       3 Assam & Meghalaya          403.
4 Bihar                      13.2       4 Bihar                      299.
5 Chhattisgarh               14.0       5 Chhattisgarh               389.
6 Coastal Andhra Pradesh      7.38      6 Coastal Andhra Pradesh     176.
```

- Removed na values

- New Columns generated and Merged data to get Final Dataset

	States.UTs	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	FirstQuarter	SecondQuarter	ThirdQuarter	FourthQuarter	ANNUAL
1	Andaman & Nicobar Islands	1901	49.20	87.10	29.20	2.30	528.80	517.50	365.10	481.10	332.60	388.50	558.20	33.60	165.50	1048.60	1178.80	980.30	3373.20
2	Andaman & Nicobar Islands	1902	0.00	159.80	12.20	0.00	446.10	537.10	228.90	753.70	666.20	197.20	359.00	160.50	172.00	983.20	1648.80	716.70	3520.70
3	Andaman & Nicobar Islands	1903	12.70	144.00	0.00	1.00	235.10	479.90	728.40	326.70	339.00	181.20	284.40	225.00	156.70	716.00	1394.10	690.60	2957.40
4	Andaman & Nicobar Islands	1904	9.40	14.70	0.00	202.40	304.50	495.10	502.00	160.10	820.40	222.20	308.70	40.10	24.10	1002.00	1482.50	571.00	3079.60
5	Andaman & Nicobar Islands	1905	1.30	0.00	3.30	26.90	279.50	628.70	368.70	330.50	297.00	260.70	25.40	344.70	4.60	935.10	996.20	630.80	2566.70
6	Andaman & Nicobar Islands	1906	36.60	0.00	0.00	0.00	556.10	733.30	247.70	320.50	164.30	267.80	128.90	79.20	36.60	1289.40	732.50	475.90	2534.40
7	Andaman & Nicobar Islands	1907	110.70	0.00	113.30	21.60	616.30	305.20	443.90	377.60	200.40	264.40	648.90	245.60	224.00	943.10	1021.90	1158.90	3347.90
8	Andaman & Nicobar Islands	1908	20.90	85.10	0.00	29.00	562.00	693.60	481.40	699.90	428.80	170.70	208.10	196.90	106.00	1284.60	1610.10	575.70	3576.40
9	Andaman & Nicobar Islands	1910	26.60	22.70	206.30	89.30	224.50	472.70	264.30	337.40	626.60	208.20	267.30	153.50	255.60	786.50	1228.30	629.00	2899.40
10	Andaman & Nicobar Islands	1911	0.00	8.40	0.00	122.50	327.30	649.00	253.00	187.10	464.50	333.80	94.50	247.10	8.40	1098.80	904.60	675.40	2687.20
11	Andaman & Nicobar Islands	1912	583.70	0.80	0.00	21.90	140.70	549.80	468.90	370.30	386.20	318.70	117.20	2.30	584.50	712.40	1225.40	438.20	2960.50
12	Andaman & Nicobar Islands	1913	84.80	0.50	1.30	2.50	190.70	530.00	280.80	205.80	580.10	288.80	133.00	67.50	86.60	723.20	1066.70	489.30	2365.80
13	Andaman & Nicobar Islands	1914	0.00	0.00	0.00	37.70	298.80	383.30	792.80	520.50	310.80	139.80	184.40	289.70	0.00	719.80	1624.10	613.90	2957.80
14	Andaman & Nicobar Islands	1915	45.00	56.70	33.30	40.90	170.20	334.70	269.00	317.20	429.80	468.10	258.40	318.00	135.00	545.80	1016.00	1044.50	2741.30
15	Andaman & Nicobar Islands	1916	0.00	0.00	0.00	0.50	487.40	450.10	317.30	425.00	561.20	369.70	192.60	133.70	0.00	938.00	1303.50	696.00	2937.50
16	Andaman & Nicobar Islands	1917	8.00	3.60	112.00	4.50	295.90	301.10	394.80	437.40	471.80	238.10	108.30	236.90	123.60	601.50	1304.00	583.30	2612.40
17	Andaman & Nicobar Islands	1918	77.40	6.90	11.40	10.70	729.30	710.80	200.90	455.40	303.30	227.00	366.90	175.00	95.70	1450.80	959.60	768.90	3275.00
18	Andaman & Nicobar Islands	1919	10.20	18.00	0.00	35.50	283.90	542.50	246.50	259.80	170.70	186.20	340.40	258.40	28.20	861.90	677.00	785.00	2352.10
19	Andaman & Nicobar Islands	1920	122.30	7.40	3.10	13.00	237.40	546.90	294.40	467.40	505.40	397.50	262.90	85.50	132.80	797.30	1267.20	745.90	2943.20
20	Andaman & Nicobar Islands	1921	13.20	3.10	0.00	37.50	351.20	282.70	487.10	330.00	581.20	360.70	118.20	41.50	16.30	671.40	1398.30	520.40	2606.40
21	Andaman & Nicobar Islands	1922	245.30	34.30	15.60	323.10	289.70	506.10	425.80	307.40	511.70	162.00	541.00	192.20	295.20	1118.90	1244.90	895.20	3554.20
22	Andaman & Nicobar Islands	1923	79.50	0.00	31.59	91.30	293.50	808.40	636.90	182.20	560.50	131.90	197.40	70.60	111.09	1193.20	1379.60	399.90	3083.79
23	Andaman & Nicobar Islands	1924	28.70	0.00	14.80	89.70	191.20	261.20	493.30	290.90	251.20	331.10	378.60	157.47	43.50	542.10	1035.40	867.17	2488.17
24	Andaman & Nicobar Islands	1925	36.60	0.00	8.60	50.40	282.20	663.80	241.80	278.20	201.90	249.50	271.50	196.00	45.20	996.40	721.90	717.00	2480.50
25	Andaman & Nicobar Islands	1926	122.10	0.00	0.00	0.50	198.40	370.00	195.30	523.70	719.30	443.80	148.40	560.70	122.10	568.90	1438.30	1152.90	3282.20
26	Andaman & Nicobar Islands	1927	3.00	17.50	17.80	108.60	504.10	433.30	195.20	370.10	126.20	327.50	274.10	65.50	38.30	1046.00	691.50	667.10	2442.90

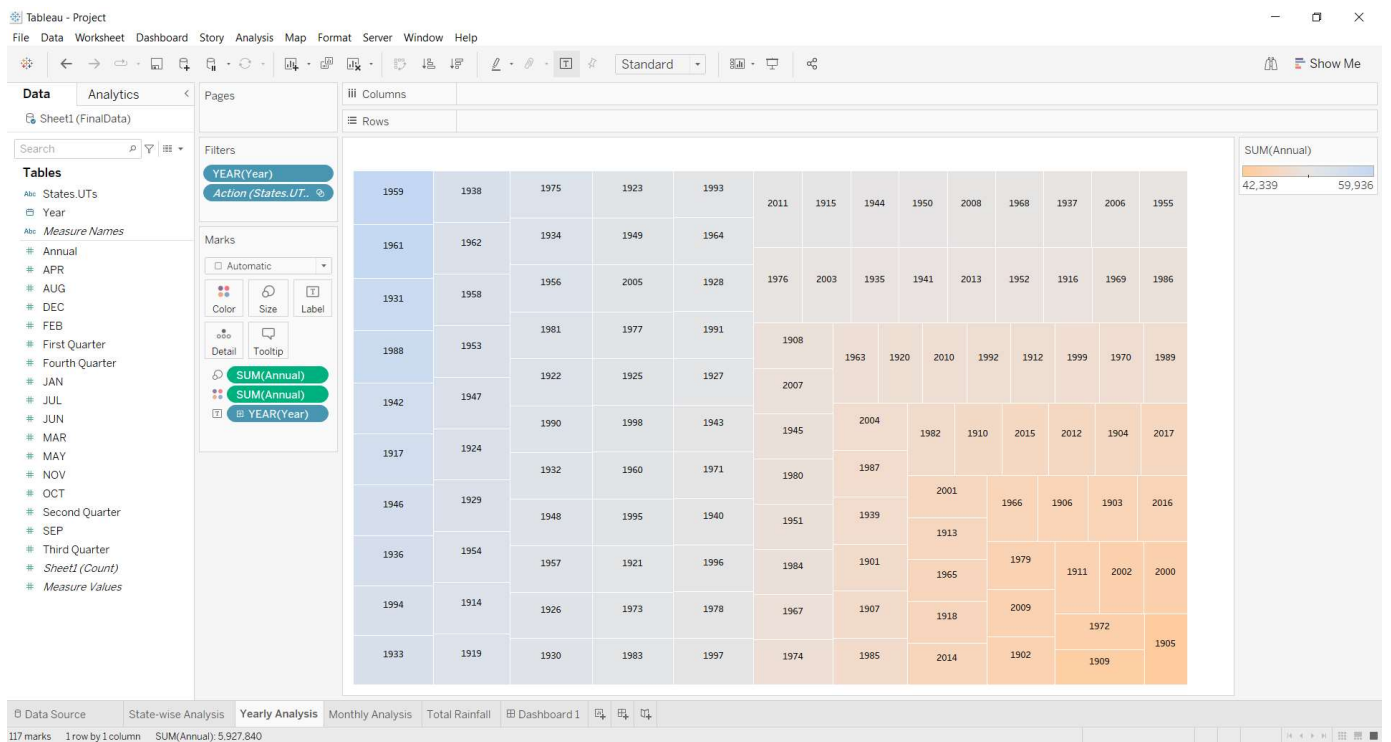Showing 1 to 27 of 4,187 entries, 19 total columns

- Tableau queries/sheets designed to prepare the Dashboard
- State-wise Analysis (Sheet 1)
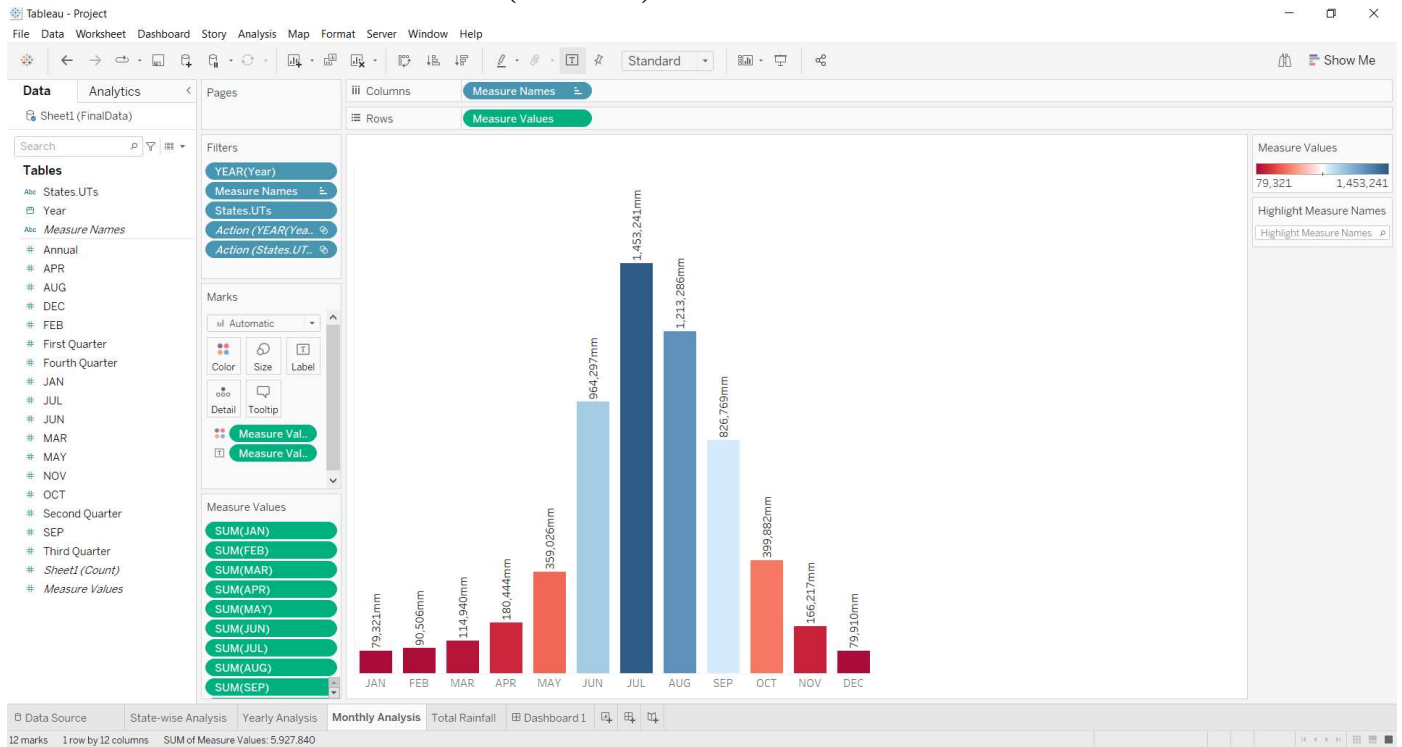
States/UTs vs Annual Rainfall
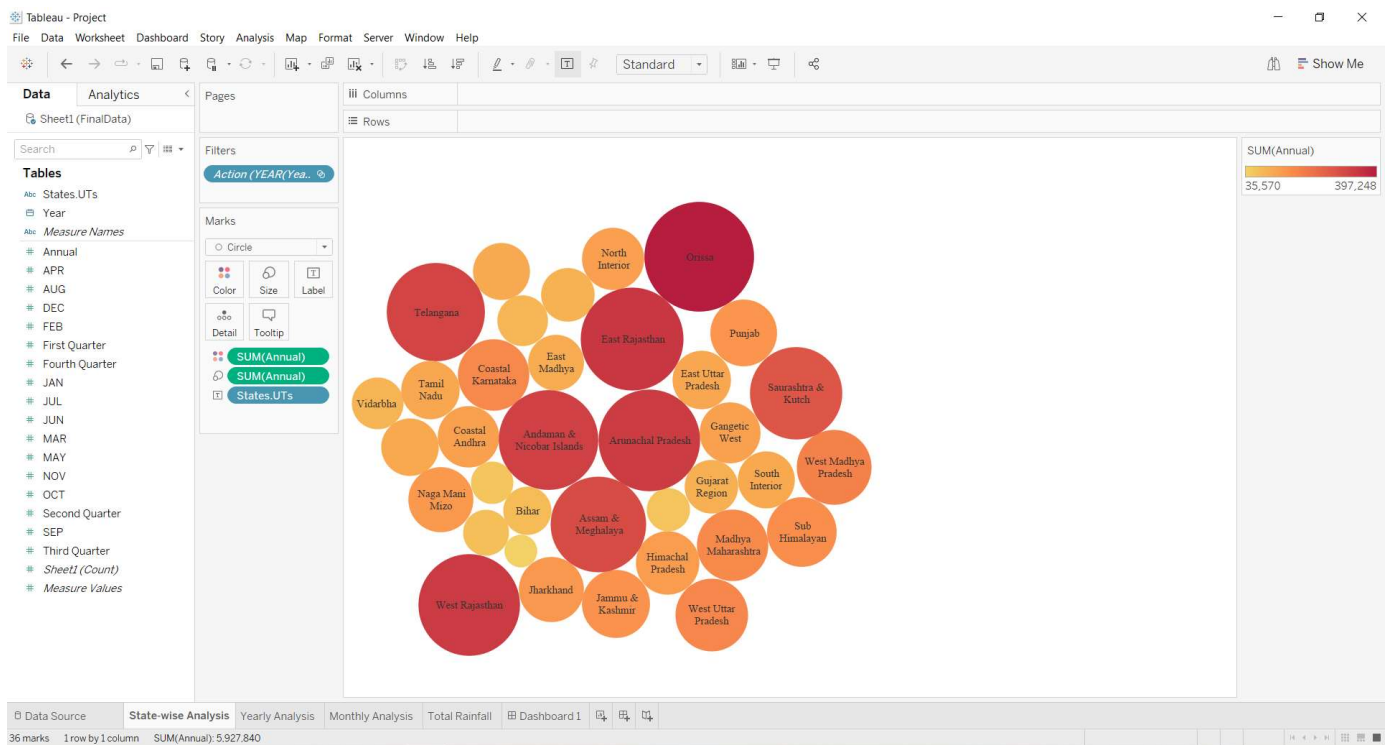
- Yearly Analysis (Sheet 2)
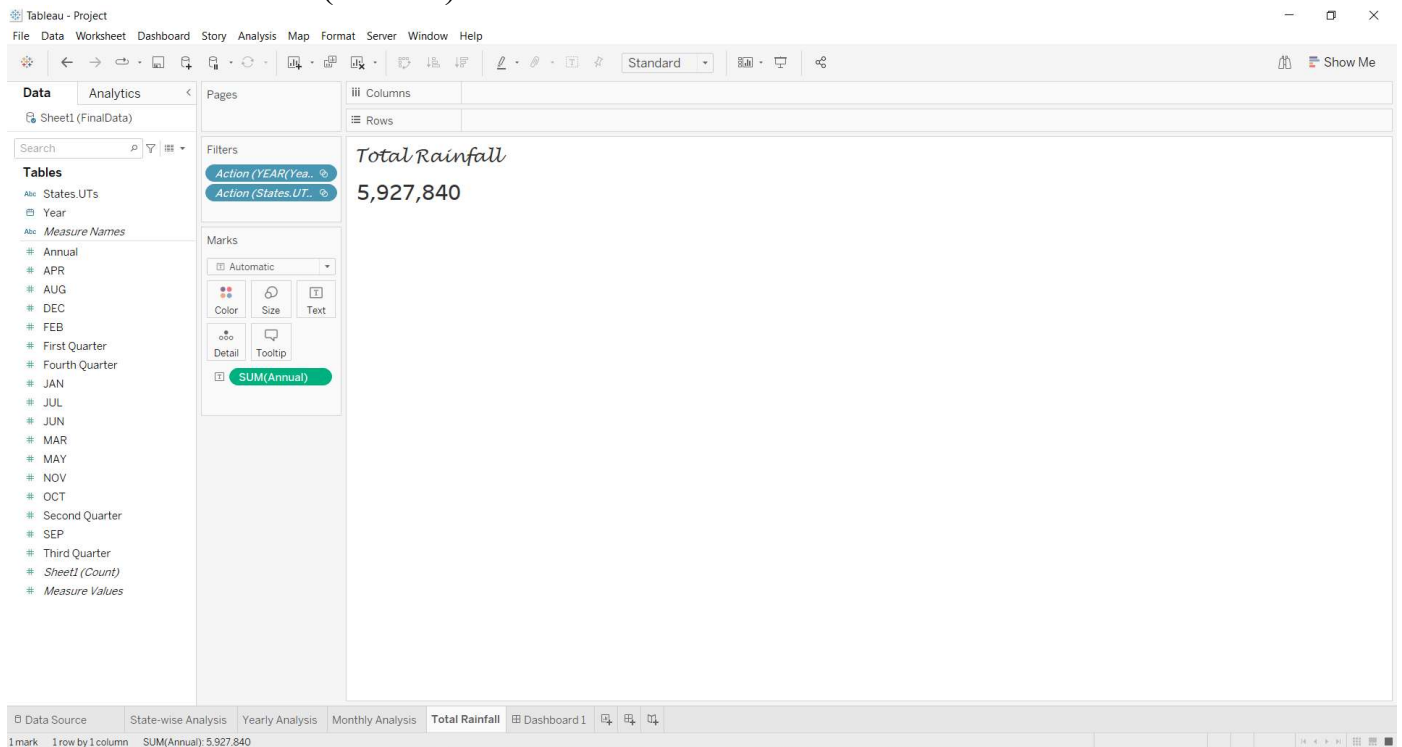
Years vs Annual Rainfall

- Monthly Analysis (Sheet 3)
  Months(Jan-Dec) vs Rainfall Value

- Total Rainfall (Sheet 4)