**Comparison of Temperature of Major Cities of India in 2022**

1. **Data Acquisition**
2. **Data Pre-processing**
3. **Statistical Analysis**
4. **Plotting Data**
5. **Interpreting Graphs**
6. **Conclusion**

**Data Acquisition:**

The data is taken from Kaggle, Weather data Indian cities (1990 to 2022). It contains Temperature data (Minimum, Average, Maximum) in degrees Centigrade and Precipitation data in mm. Source: <https://www.kaggle.com/datasets/vanvalkenberg/historicalweatherdataforindiancities>

**Data Pre-processing:**

For our study we are going to conduct our analysis on only three cities Bangalore, Mumbai, and Delhi respectively. Moreover, we will only use the average temperature data to keep it simple.

**Statistical Analysis:**

Hypothesis that we worked on was “There’s a reasonable difference between the temperature of the three cities in year 2022”. As we've three independent variables (cities) and one dependent variable (temperature) we will be using One-Way Analysis of Variance (ANOVA) to compare the temperature among these three cities.

Data Analysis Tool pack in Microsoft Excel 365 is used for analysis of our data. ANOVA: Single Factor is used. After conducting the Analysis among these 3 groups we got the p-value p = 1.007E-28 Its less than alpha (0.05) It shows there is significant difference among the three cities.

In order to evaluate this significant difference further we applied POST-HOC Tests where we made three combinations of the cities. Bangalore vs Mumbai, Bangalore vs Delhi, and Mumbai vs Delhi. Two-tailed equal variance T-test was applied to each combination. On T-test we applied Bonferroni and got Critical-P and according to it the difference was still significant. To take it even further Tukey HSD was applied q-Tukey was calculated and compared with q-Critical and it showed the same level of significance among the differences as by Bonferroni. Thus, the temperature of each city is significantly different than the other in our dataset.

**Plotting Data:**

A bar-plot of means of temperatures was plotted using the built-in plot function of Microsoft Excel 365 along with displaying the annotations of significance.

**Interpreting Graphs:**

The graph clearly indicates the difference among the temperature of the three cities. Bangalore had 22.6 °C, Mumbai had 27.6 °C and Delhi had 31.4 °C average temperature throughout 2022. The annotated asterisks indicate the significant difference among the cities.

**Conclusion:**

Our results indicate that the hypothesis stands corrected. There is a reasonable difference of temperature among the three cities. The percentage difference of temperature between the combinations B vs M, B vs D, M vs D was 22%, 39% and 17% respectively. Results also indicate Delhi to be the hottest while Bangalore to be the coolest city in our dataset.

**References:**

Tukey Test: <https://www.automateexcel.com/stats/tukey-test/>