**REPORT ON STATISTICAL ANALYSIS OF HEALTH OUTCOMES AND ENVIRONMENTAL FACTORS ( PROJECT 1 IN COLLEGE OF SCIENCES)**

**INSTITUTION: BELLS UNIVERSSITY OF TECHNOLOGY**

**COURSE CODE: ICT 313**

**DEPARTMENT: BIOCHEMISTRY**

**GROUP NO : 3**

**GROUP MEMBERS**

1. Jimoh Zuliyah (2022/11117)
2. Arosoye Glory (2022/11726)
3. Adekemi Oluwatosin (2022/11684)
4. Nwaobi Ugochukwu (2022/11369)
5. Alabi Roseline ( 2022/11720)
6. Ikwuka Joshua chidera (2022/11651)

**TABLE OF CONTENTS**

1. Project Overview
2. Introduction
3. Aim and Objectives
4. Techniques
5. Data Description
6. Analysis and Result
7. Conclusion

**PROJECT SUMMARY**

The statistic application project requires selecting a topic aligned with departmental strength. Each group must choose a topic from the predefined section related to IBM SPSS. It is recommended to use relevant articles and journals to support the project. A report meeting academic standard and proper referencing must be submitted. All completed project must be uploaded to GitHub with all group members listed as contributors.

**INTRODUCTION**

The statistical analysis of health outcomes and environmental factors explores the relationship between environmental exposures and public health. By examining these factors through data, we can identify patterns and assess their impact on diseases, mortality rates, and overall well-being. This topic is crucial due to the growing concerns over environmental degradation and its direct and indirect effects on human health.

**AIMS AND OBJECTIVES**

The aims and objectives of statistical analysis of health outcomes and environmental factors are to:

Aims:

1. Assess environmental impact on health outcomes.
2. Identify risk factors and patterns in health.
3. Inform evidence-based policy for public health.
4. Evaluate intervention effectiveness.
5. Develop predictive models for health impacts.

Objectives:

1. Collect reliable data on health and environmental factors.
2. Apply statistical models to explore relationships.
3. Quantify associations between environmental factors and health.
4. Identify vulnerable populations at higher risk.
5. Estimate long-term health effects of environmental exposures.

**TECHNIQUES**

The analysis was performed using IBM SPSS, a statistical software designed for data analysis. The techniques supplied includes:

1. Correlation Analysis: It is used to examine the relationships between environmental factors (PM, UFP, WI, GS) and health outcomes (HFI, PCTS, PRI) using Pearson correlation.

2. Regression Analysis: It was used to test how environmental factors (PM, UFP, WI, GS) predict the Health Factor Index (HFI) using linear regression.

3. Spatial Analysis: It is used to visualize the geographic distribution of data points using East (longitude) and North (latitude) coordinates.

4. Descriptive Statistics: It is used to calculate basic statistics (mean, standard deviation, minimum, maximum) for key variables (PM, UFP, HFI, PCTS).

**DATA DESCRIPTION**

* Environmental Variables (Independent Variables):
* NEI (Neighbourhood Environmental Index): A measure of environmental quality in the neighbourhood.
* PM (Particulate Matter): Air quality indicator (e.g., PM2.5, PM10).
* UFP (Ultrafine Particles): Another air quality indicator.
* WI (Walkability Index): Measures how walkable the neighbourhood is.
* GS (Green Space): Indicates the amount of green space in the area.
* FF (Fuel Consumption): Could be related to air pollution.
* HP (Heat Pollution): Indicates urban heat island effects.
* SN (Noise Pollution): Environmental noise levels.
* Health Outcome Variables (Dependent Variables):
* HFI (Health Factor Index): A composite measure of health outcomes.
* PCTS (Prevalence of Chronic Conditions): Indicates the prevalence of chronic diseases.
* PNCTS (Prevalence of Non-Chronic Conditions): Indicates the prevalence of non-chronic diseases.
* PRI (Premature Mortality Rate): Measures premature deaths.
* TC (Total Cases): Could represent the total number of health-related cases.

**ANALYSIS AND RESULT**

**Analysis 1:** Pearson's Correlation Coefficient

If PM and PCTS have an r = 0.75 with p < 0.05, it means there’s a strong positive correlation between air pollution and chronic disease prevalence.

Correlation Coefficients (r):

Look under "Pearson Correlation" for each pair of variables.

Example: PM and UFP have r = -0.170, indicating a weak negative correlation.

Significance (p-values):

Look under "Sig. (2-tailed)".

Example: UFP and PRI have r = 0.578 with p < 0.001, indicating a strong positive correlation that is statistically significant.

Key Findings:

UFP and PRI are strongly correlated (r = 0.578, p < 0.001)

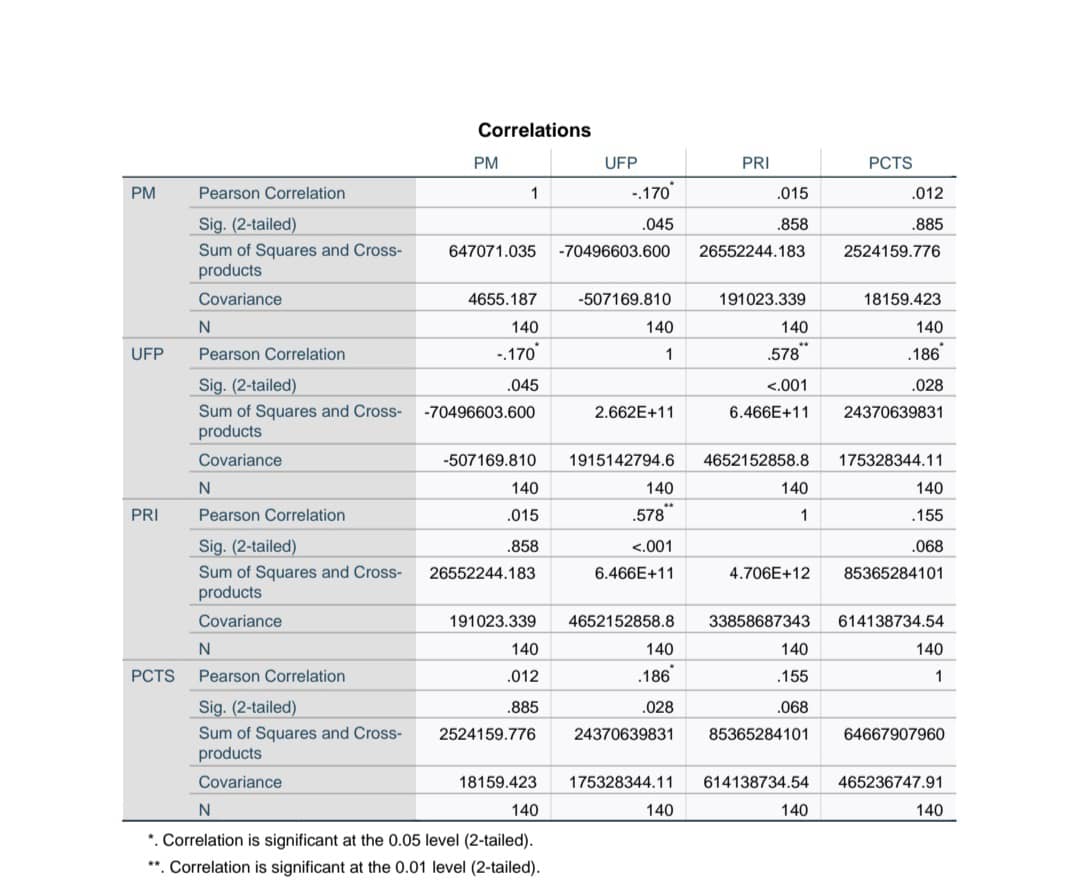
Other correlations (e.g., PM with PCTS) are weak and not significant (p > 0.05).

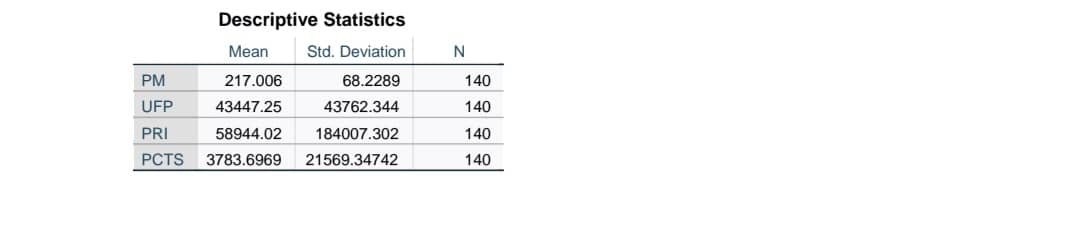
N: Sample size is 140 for all variables.

Summary:

Strongest correlation: UFP and PRI (r = 0.578, significant).

Weak/no correlation: Most other pairs (e.g., PM with PCTS)



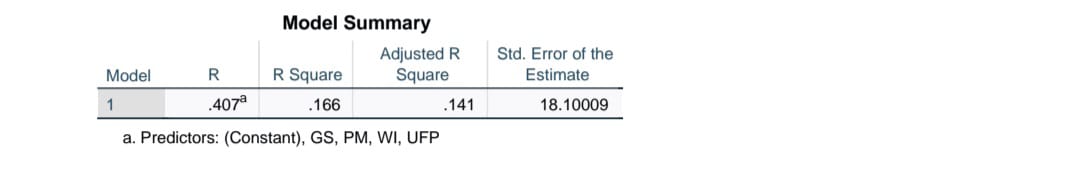


**Analysis 2** : Regression analysis

1. Model Summary Table

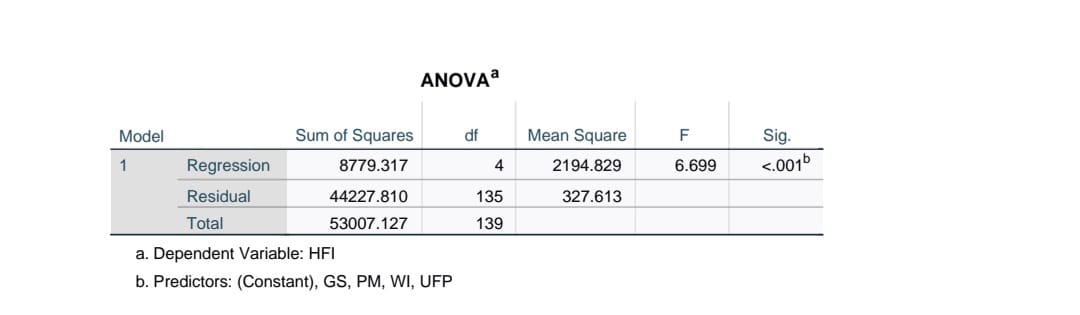
R² = 0.166: The model explains 16.6% of the variance in the dependent variable (HFI).

Adjusted R² = 0.141: Adjusted for the number of predictors, the model explains 14.1% of the variance.



2. ANOVA Table

F = 6.699, p < 0.001: The overall regression model is statistically significant, meaning at least one predictor significantly predicts HFI.



3. Coefficients Table

Significant Predictors:

WI (Walkability Index):

B = 0.568, p < 0.001: A 1-unit increase in WI is associated with a 0.568-unit increase in HFI, and this effect is statistically significant.

Non-Significant Predictors:

PM (Particulate Matter): p = 0.500 (not significant).

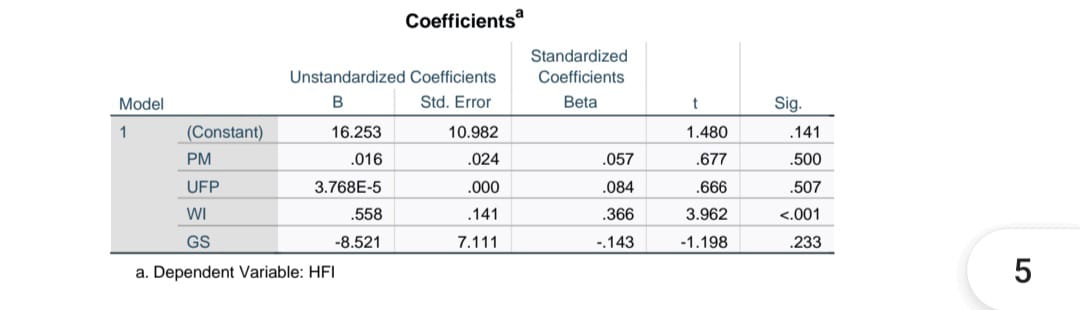
UFP (Ultrafine Particles): p = 0.507 (not significant).

GS (Green Space): p = 0.233 (not significant).

Key Inference:

Walkability Index (WI) is the only significant predictor of HFI in this model.

The model explains a moderate amount of variance (16.6%) in HFI, but other environmental factors (PM, UFP, GS) do not significantly contribute to predicting HFI.



**Analysis 3:** Spatial analysis: Results:

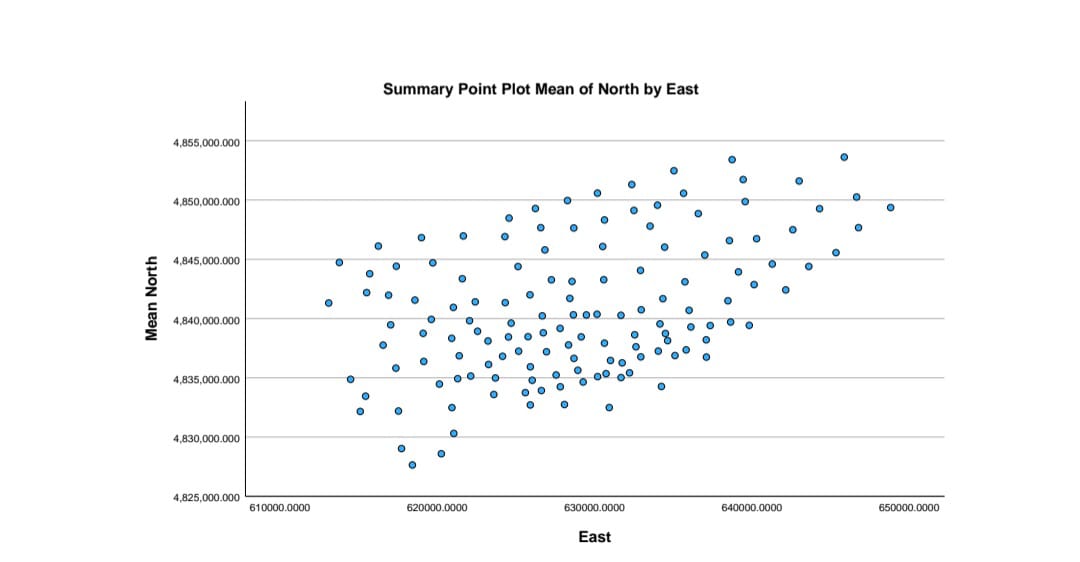
The plot displays the geographic distribution of data points across the East (longitude) and North (latitude) coordinates.

Each point represents the mean North value for a given East value.

Inference:

The points are spread across a range of East values (610,000 to 650,000) and North values (4,825,000 to 4,855,000).

There is no clear clustering or pattern in the distribution, suggesting no strong spatial trend in the data based on the variables plotted.



**Analysis 4:** Descriptive Statistics:

Results:

PM (Particulate Matter):

Mean = 217.006, Std. Deviation = 68.2289: Moderate variability in air pollution levels.

UFP (Ultrafine Particles):

Mean = 43,447.25, Std. Deviation = 43,762.344: High variability in ultrafine particle levels.

HFI (Health Factor Index):

Mean = 60.1909, Std. Deviation = 19.52809: Moderate variability in health outcomes.

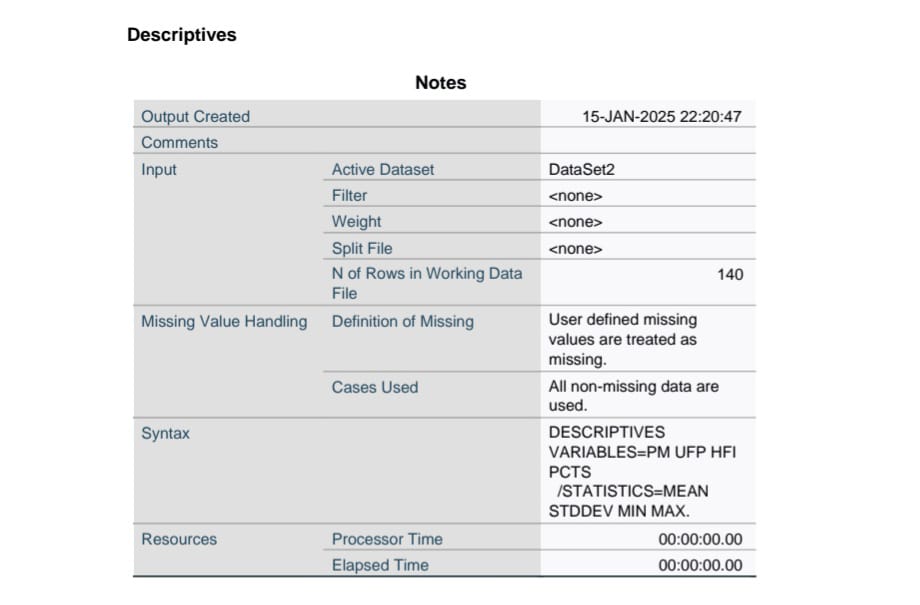
PCTS (Prevalence of Chronic Conditions):

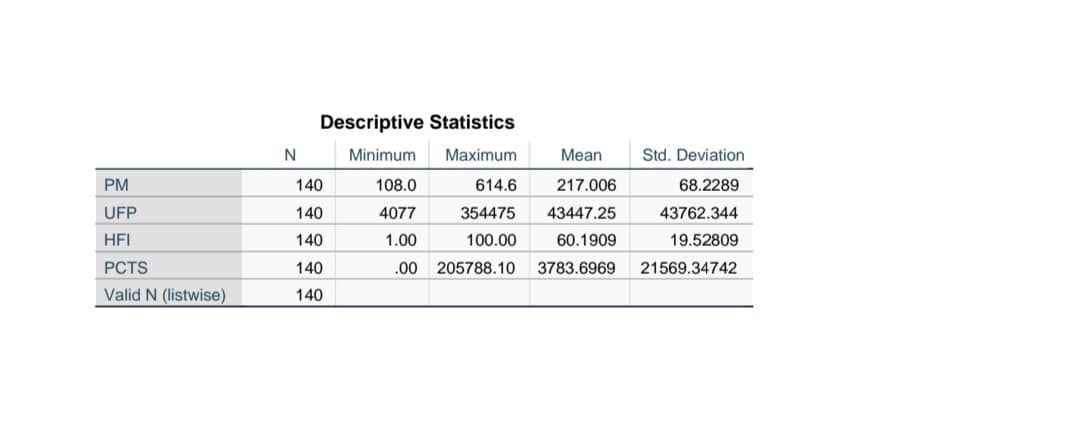
Mean = 3,783.6969, Std. Deviation = 21,669.34742: High variability in chronic disease prevalence.

Inference:

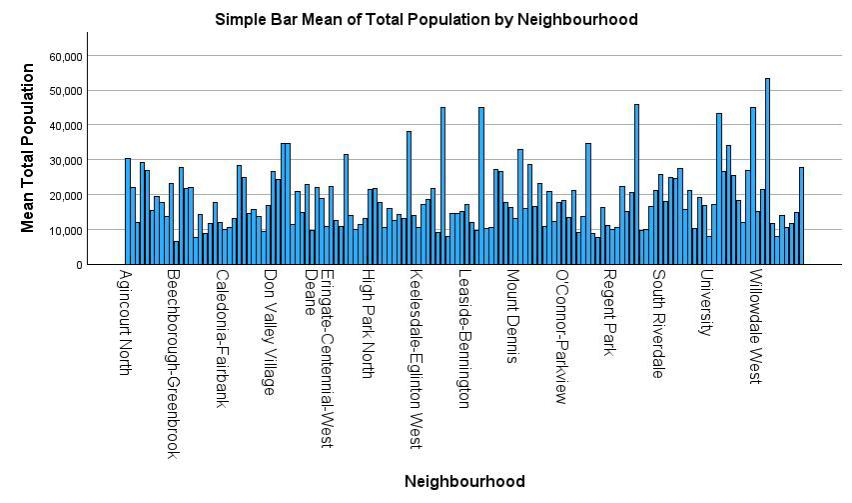
PM and HFI show moderate variability, while UFP and PCTS have high variability.

The wide range in UFP and PCTS suggests significant differences across neighbourhoods.





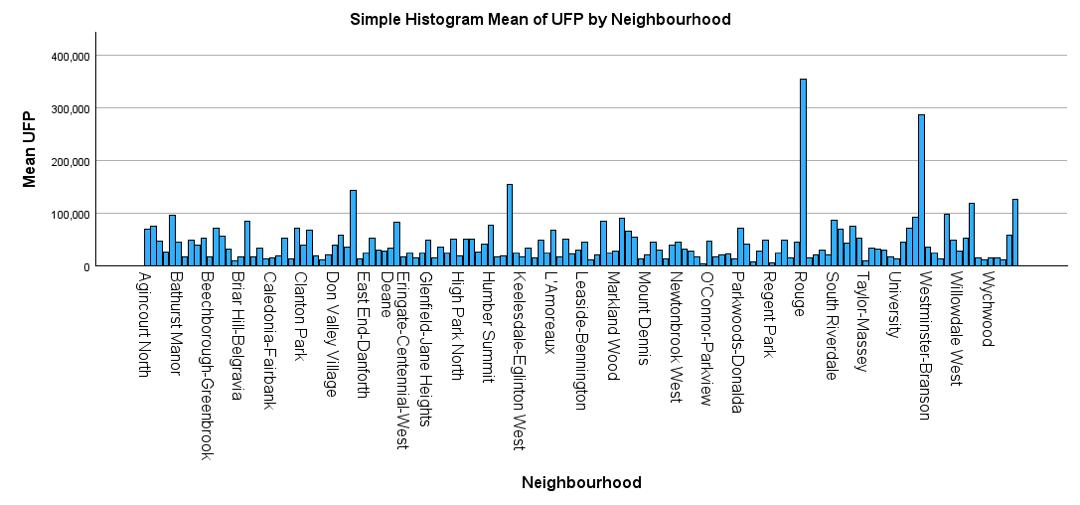
**VISUALISATIONS**

****

**BAR CHART**

Summary:

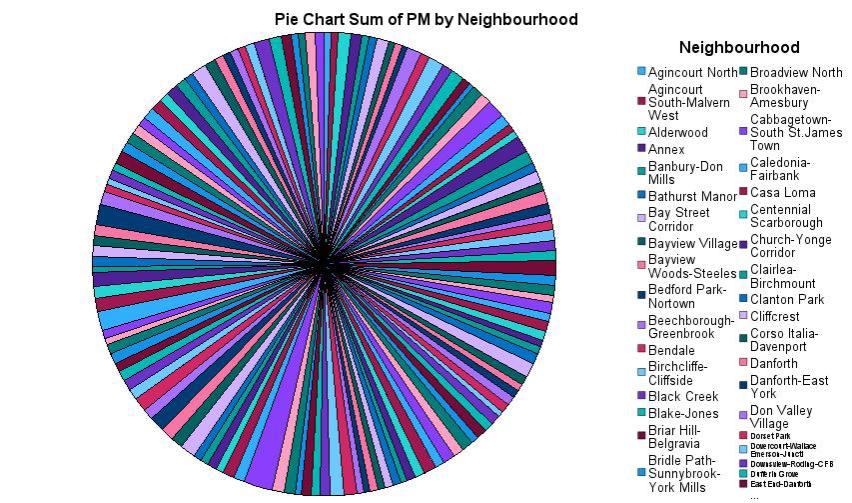
The bar chart reveals significant differences in population sizes across neighborhoods, highlighting variations in urbanization and density. Neighborhoods with larger populations may require more resources and infrastructure, while those with smaller populations might have different community characteristics. This data can inform urban planning and resource allocation strategies.



**HISTOGRAM**

Summary:

The histogram reveals significant differences in UFP levels across neighborhoods, highlighting potential air quality and health disparities. Neighborhoods with higher UFP levels may require targeted interventions to improve environmental conditions and public health outcomes.



PIE CHART

**CONCLUSION**

In conclusion, the statistical analysis demonstrates a significant relationship between environmental factors and health outcomes. Variations in factors such as air quality, water access, and climate conditions are associated with changes in health metrics, including disease prevalence and overall well-being. These findings highlight the importance of addressing environmental factors to improve public health and guide policy interventions aimed at mitigating negative health impacts linked to environmental conditions.