## ****1. Advantages of Analytics and Business Intelligence for Decision-Making (Pyramid Structure)****

The integration of analytics and business intelligence (BI) in decision-making processes provides significant advantages across the **strategic**, **tactical**, and **operational** levels of management. These insights empower organizations, including ministries like Sri Lanka's Ministry of Industry and Commerce, to enhance decision-making efficiency, precision, and alignment with national goals.

## ****1.1 Strategic Level****

At the strategic level, analytics and BI focus on long-term planning, policy-making, and resource allocation. These tools allow senior executives to analyze trends, predict future scenarios, and evaluate the impact of decisions on the national economy.

* **Economic Trend Forecasting:**
  + Using predictive analytics, governments can identify high-potential industries for future investment. For instance, Sri Lanka’s Export Development Board utilizes BI dashboards to prioritize growth in sectors like tourism and IT (World Bank, 2020).
  + Similarly, **Canada** employs labor market analytics to forecast occupational demand and align education policies with workforce needs (OECD, 2021).
* **Policy Impact Simulation:**
  + Analytics tools allow policymakers to assess the potential socio-economic outcomes of policy changes. For example, **Singapore’s Smart Nation Initiative** evaluates urban development policies to ensure long-term sustainability and economic resilience (Smart Nation Singapore, 2020).
* **National Development Prioritization:**
  + By analyzing global market trends, governments can identify export opportunities. For instance, **India’s Skill Development Mission** used analytics to identify high-growth industries such as IT and renewable energy, steering education and training programs accordingly (Government of India, 2019).

## ****1.2 Tactical Level****

Middle management relies on BI and analytics to translate strategic goals into actionable plans. This involves resource allocation, performance monitoring, and scenario analysis.

* **Resource Optimization:**
  + BI tools provide insights into optimal allocation of resources. In **South Korea**, analytics guided investments in industrial zones, ensuring balanced development across regions (Korea Institute for Industrial Economics, 2021).
* **KPI Monitoring and Adjustments:**
  + Dashboards help monitor project performance metrics in real-time. For instance, Sri Lanka’s garment sector utilizes BI to track production efficiency and export targets (FAO, 2021).
* **Workforce Scenario Modeling:**
  + Scenario analysis tools support workforce planning under uncertain conditions. **Australia’s Industry Growth Centers** use data to align workforce skills with the needs of growing sectors such as construction and healthcare (Australian Government, 2020).

## ****1.3 Operational Level****

At the operational level, frontline managers use analytics to enhance daily activities, ensuring alignment with broader organizational goals.

* **Automation of Reports:**
  + Automated reporting reduces manual effort and errors. For example, Sri Lanka’s Customs Department leverages BI dashboards to streamline trade documentation processes (ADB, 2020).
* **Workforce Scheduling:**
  + Analytics optimizes labor deployment in real-time. In **Bangladesh**, garment factories use workforce analytics to improve scheduling efficiency and minimize costs (ILO, 2020).
* **Real-Time Decision Support:**
  + BI tools assist in addressing operational challenges. The **UK’s National Health Service (NHS)** predicts hospital resource requirements using analytics, ensuring timely care delivery (NHS Digital, 2021).

## ****2. Tools, Techniques, and Methodologies for Case Study Analysis****

The aim of this project is to identify factors influencing occupational prestige and provide actionable insights to the Ministry of Industry and Commerce in Sri Lanka. A variety of tools, techniques, and methodologies are used to ensure accurate and efficient analysis.

## ****2.1 Tools****

**Microsoft Excel:**

* Used for data cleaning, initial exploration, and summary statistics.
* It provides quick insights into measures of central tendency (mean, median, mode) and variability (standard deviation, range).
* Suitable for generating basic charts such as histograms and scatter plots.

**RStudio (R Programming):**

* A powerful statistical software for advanced data analysis.
* Used to perform regression analysis, correlation studies, and hypothesis testing.
* Ideal for generating complex visualizations like density plots and cluster diagrams.

**SPSS (Statistical Package for the Social Sciences):**

* Useful for conducting robust statistical tests and descriptive analyses.
* Supports ANOVA, t-tests, and other methods to assess the significance of relationships between variables.

**Tableau/Power BI:**

* Visualization tools for presenting data in an interactive and intuitive way.
* Used to create dashboards summarizing prestige, education, and income trends with dynamic filtering options.

## ****2.2 Techniques****

* **Data Cleaning and Preprocessing:**
  + Preparing the dataset by handling missing values, removing outliers, and standardizing numerical variables. This ensures accurate analysis.
  + Example: Normalizing income and education data for consistent comparison.
* **Descriptive Statistics:**
  + Summarizing the dataset through measures like mean, median, and standard deviation to understand the general behavior of the data.
  + Example: Calculating the average prestige score across different occupation types.
* **Correlation Analysis:**
  + Measuring the strength of relationships between variables such as education, income, and prestige.
  + Example: Identifying a strong positive correlation between higher education levels and occupational prestige.
* **Regression Analysis:**
  + Building models to predict prestige scores based on independent variables like income, gender representation, and education.
  + Example: A regression model explaining how income and gender equality contribute to prestige variation.
* **Clustering:**
  + Grouping occupations into clusters based on similarities in factors like income and education. This technique helps identify patterns in the dataset.
  + Example: Grouping occupations into clusters with high, medium, and low prestige scores.

## ****2.3 Methodologies****

* **Exploratory Data Analysis (EDA):**
  + Investigating data patterns, trends, and relationships to form hypotheses about what influences prestige.
  + Example: EDA reveals that occupations with higher female participation often have varying prestige scores.
* **Hypothesis Testing:**
  + Formulating hypotheses about factors influencing prestige and testing their validity through statistical methods.
  + Example: Testing if education level significantly impacts occupational prestige.
* **Predictive Modeling:**
  + Creating models to predict prestige scores based on variables like income and education.
  + Example: A linear regression model predicts how a 10% increase in income affects prestige.
* **Global Comparisons:**
  + Comparing results with international datasets to validate insights and provide a broader perspective.
  + Example: Comparing factors influencing prestige in Canada with those in Sri Lanka to identify similarities and differences.

## ****2.4 Summary of This Project****

The project focuses on analyzing a dataset related to occupations in Canada to determine the factors that influence occupational prestige. The insights will help the Ministry of Industry and Commerce in Sri Lanka make informed decisions for national development. The key steps include:

**Data Cleaning:** Preparing the dataset by removing inconsistencies and ensuring it is ready for analysis.

**Exploratory Analysis:** Identifying trends and relationships between factors like income, education, and prestige.

**Statistical Analysis:** Applying regression and correlation analysis to uncover key drivers of occupational prestige.

**Model Building:** Creating predictive models to forecast how changes in specific variables affect prestige.

**Visualization:** Using tools like Tableau to create dashboards for presenting findings.

**Reporting:** Providing a comprehensive report with actionable insights tailored to the ministry's objectives.

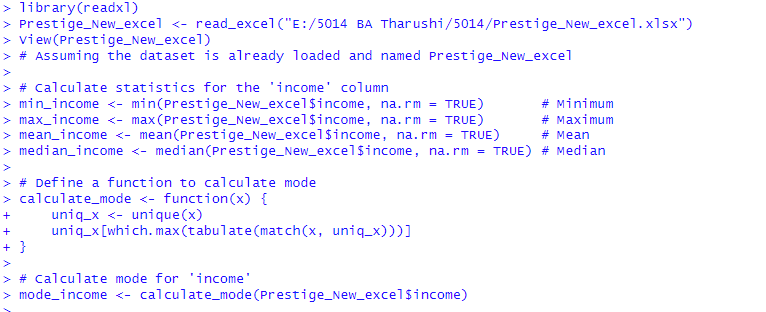
This structured approach ensures that the ministry can leverage data science to prioritize occupations critical to Sri Lanka’s national development, aligning with global best practices.

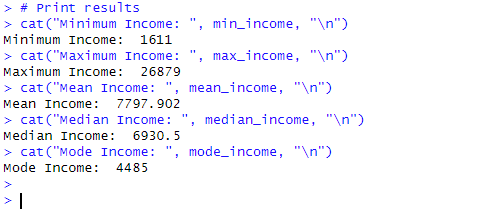
## ****3.1 Summary Statistics for Income****

To analyze the distribution of income among incumbents, the following statistical measures were calculated:

* **Minimum Income:**
  + The lowest recorded income is **1,611**.
  + This indicates the baseline income within the dataset and highlights the range of lower-income occupations.
* **Maximum Income:**
  + The highest recorded income is **26,879**.
  + This represents the upper threshold of income and typically correlates with high-prestige occupations requiring advanced skills or education.
* **Mean Income (Average):**
  + The average income across all incumbents is approximately **7,797.90**.
  + This provides a central tendency measure, showing the typical income level.
* **Median Income:**
  + The median income is **6,930.50**.
  + The median, being slightly lower than the mean, indicates a right-skewed distribution where a few higher incomes elevate the mean.
* **Mode Income:**
  + The most frequently occurring income is **4,485**.
  + This suggests that a significant number of occupations fall within this income range.

## 3.2 R studio Codes for Solution





## ****3.2 Interpretation****

The calculated measures give a clear understanding of the income distribution:

The **minimum and maximum values** highlight the broad range of incomes, reflecting the diversity of occupations in the dataset.

The **mean and median values** indicate that income is not symmetrically distributed, with a skew towards higher values due to a few high-income occupations.

The **mode** shows the most common income level, providing insights into the concentration of incomes.

These statistics are essential for understanding how income influences the prestige of occupations. Occupations with higher income typically align with higher prestige, making income a critical factor in decision-making for resource allocation and workforce planning.

## ****4.1 Summary Statistics and Descriptive Justifications****

The summary statistics for **prestige**, **education**, and **income** are calculated to understand the central tendency, dispersion, and range of values. Here's an analysis followed by descriptive justifications:

## ****4.1.1 Prestige****

**Mean:** 56.83

* The average prestige score of occupations is moderately high, indicating that most occupations in the dataset are associated with respectable social status.

**Standard Deviation (Std):** 17.20

* This suggests a considerable variation in prestige scores, showing diversity in how different occupations are perceived socially.

**Minimum and Maximum:** 24.8 and 89.5

* The range indicates that some occupations are considered low in prestige, while others are highly prestigious.

**Mode:** 50.6

* The most frequently observed prestige score aligns closely with the mean, indicating a concentration around the middle range.

**Justification:** The variation in prestige scores suggests a diverse representation of occupations in the dataset, from low-prestige manual labor jobs to high-prestige professions such as medicine or law. These scores provide critical insights into occupational stratification and societal valuation.

## ****4.1.2 Education****

**Mean:** 10.74

* On average, incumbents have around 10.7 years of education, reflecting a moderately educated workforce.

**Standard Deviation (Std):** 2.73

* Education levels vary significantly, highlighting differences in qualifications required for various occupations.

**Minimum and Maximum:** 6.38 and 16.4

* The minimum indicates jobs requiring only basic education, while the maximum corresponds to highly specialized fields requiring extensive education.

**Mode:** 9.5

* The most common education level is slightly below the mean, indicating a significant number of mid-education-level occupations.

**Justification:** Education is a crucial factor influencing occupational prestige. Higher education levels often correlate with higher prestige, reflecting societal emphasis on academic qualifications for professional success.

## ****4.1.3 Income****

**Mean:** 7,797.90

* The average income is relatively moderate, with some occupations earning significantly more than others.

**Standard Deviation (Std):** 4,245.92

* A high standard deviation suggests substantial income disparity among occupations.

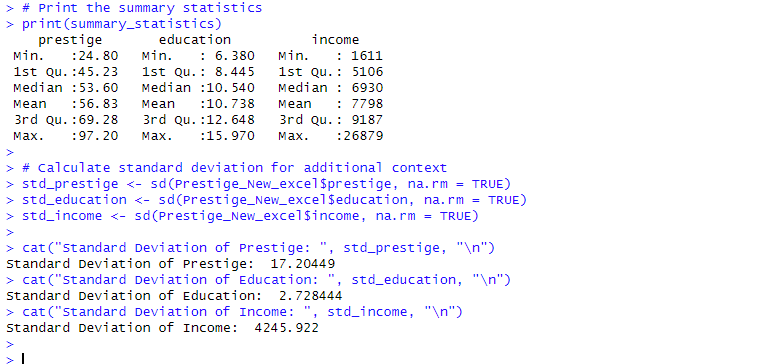
**Minimum and Maximum:** 1,611 and 26,879

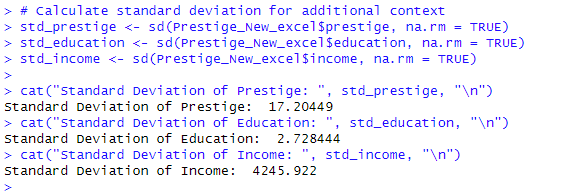
* This wide range shows the economic inequality between low-paying jobs and high-paying prestigious professions.

**Mode:** 4,485

* The most frequently occurring income is lower than the mean, indicating that many occupations fall in the lower-income bracket.

**Justification:** Income plays a vital role in determining the prestige of an occupation. High-income jobs tend to align with higher prestige, underscoring the economic dimensions of social valuation.





## ****5.1 Central Tendency Analysis for Prestige, Education, and Income****

Central tendency analysis focuses on summarizing data around central points using measures such as **mean**, **median**, and **mode**. This analysis evaluates the **prestige**, **education**, and **income** of incumbents, providing insights into their distribution and variability. Bell curves (density plots) visually represent these distributions, making the patterns and tendencies clearer.

## ****5.1.1 Prestige****

**Mean:** 56.83

**Median:** 53.6

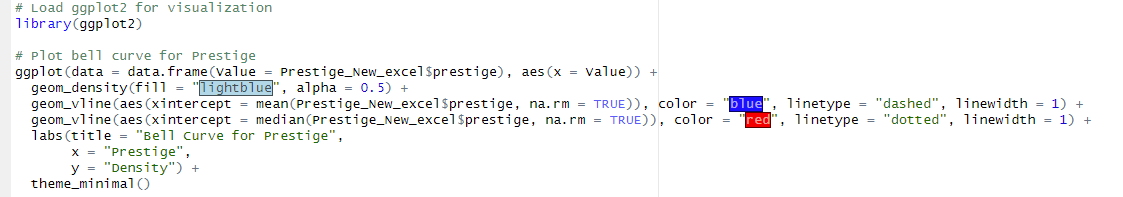
**Mode:** 61.1

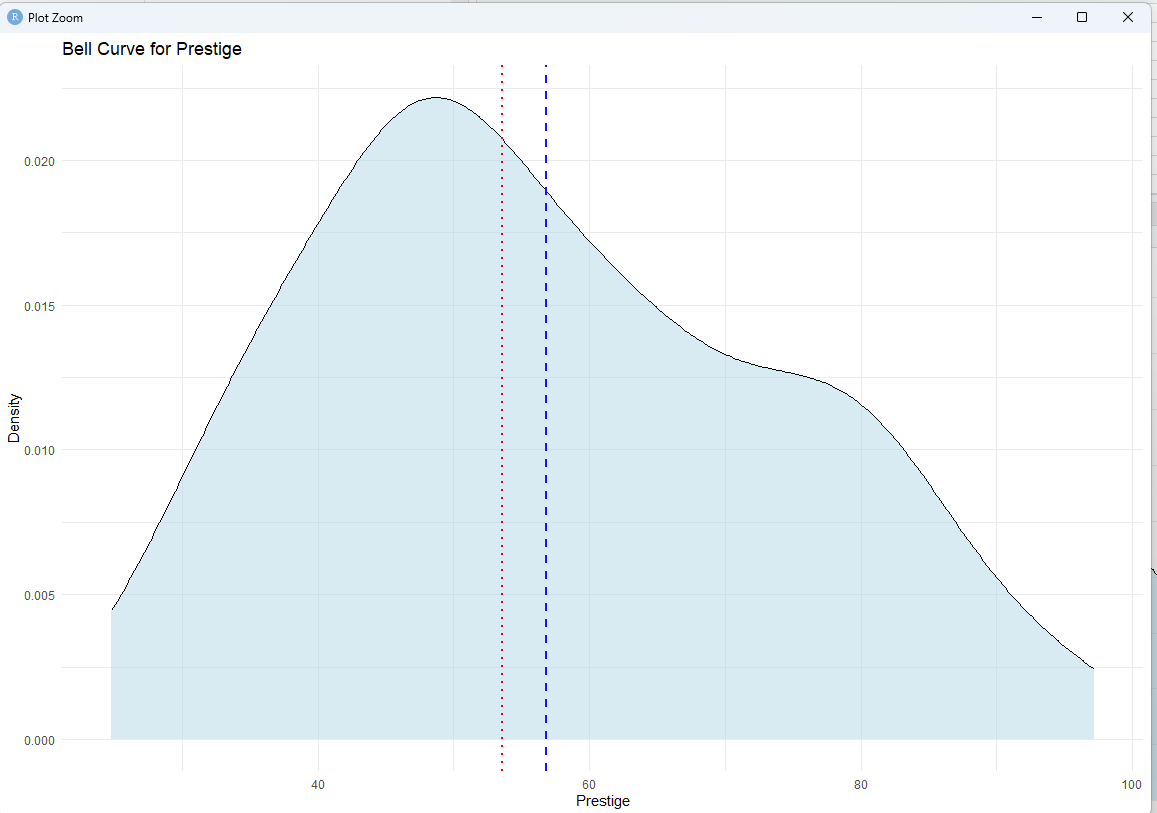
**Standard Deviation:** 17.20

**Interpretation:**  
The mean prestige score indicates that occupations, on average, hold moderate societal respect. The proximity of the mean and median suggests that the distribution is relatively symmetric. However, the slight difference points to a minor skew caused by outlier values at both ends of the scale. The mode (61.1) shows the most frequently occurring prestige score. The bell curve highlights a moderate spread, reflecting diverse social valuation across occupations, ranging from low-prestige roles to highly regarded professions.

**Justification:**  
The moderate standard deviation suggests that while most occupations cluster around the mean, there is considerable variability. This variability aligns with the presence of both low-prestige and high-prestige jobs in the dataset, such as manual labor roles versus professional positions like medicine or engineering.

## 5.1.2 Bell Curve for Prestige





## ****5.2 Education****

**Mean:** 10.74 years

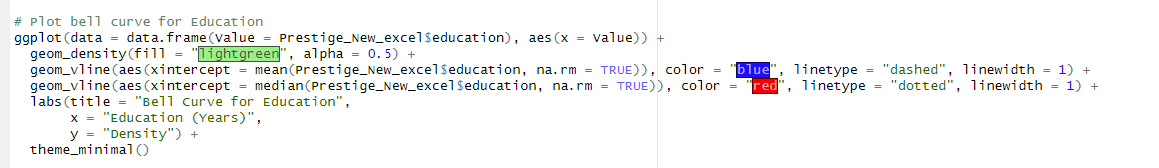
**Median:** 10.54 years

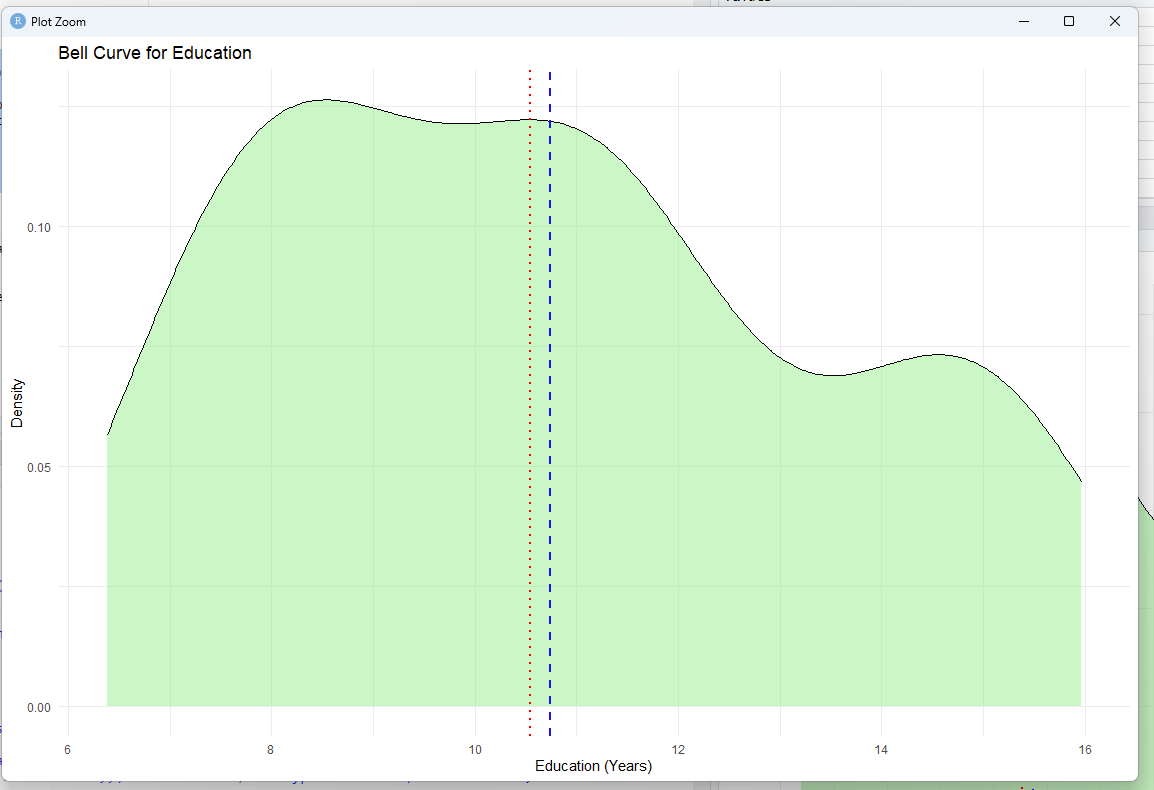
**Mode:** 13.62 years

**Standard Deviation:** 2.73

**Interpretation:**  
The mean and median years of education required suggest that most occupations demand a moderately high level of formal education. The mode (13.62 years) represents jobs requiring higher educational qualifications, like specialized professions. The bell curve for education is relatively narrow, indicating that education levels among occupations are less varied compared to prestige and income.

**Justification:**  
The narrow distribution reflects the structured nature of educational requirements, as most occupations fall within a range of 8–14 years of education. Higher education levels often correlate with higher prestige, demonstrating the importance of academic qualifications in workforce planning.





## ****5.3 Income****

**Mean:** 7,798

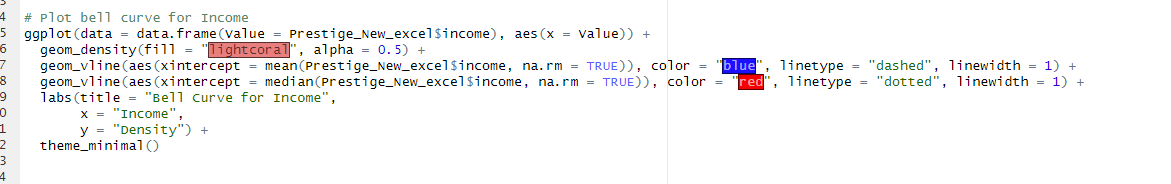
**Median:** 6,930.5

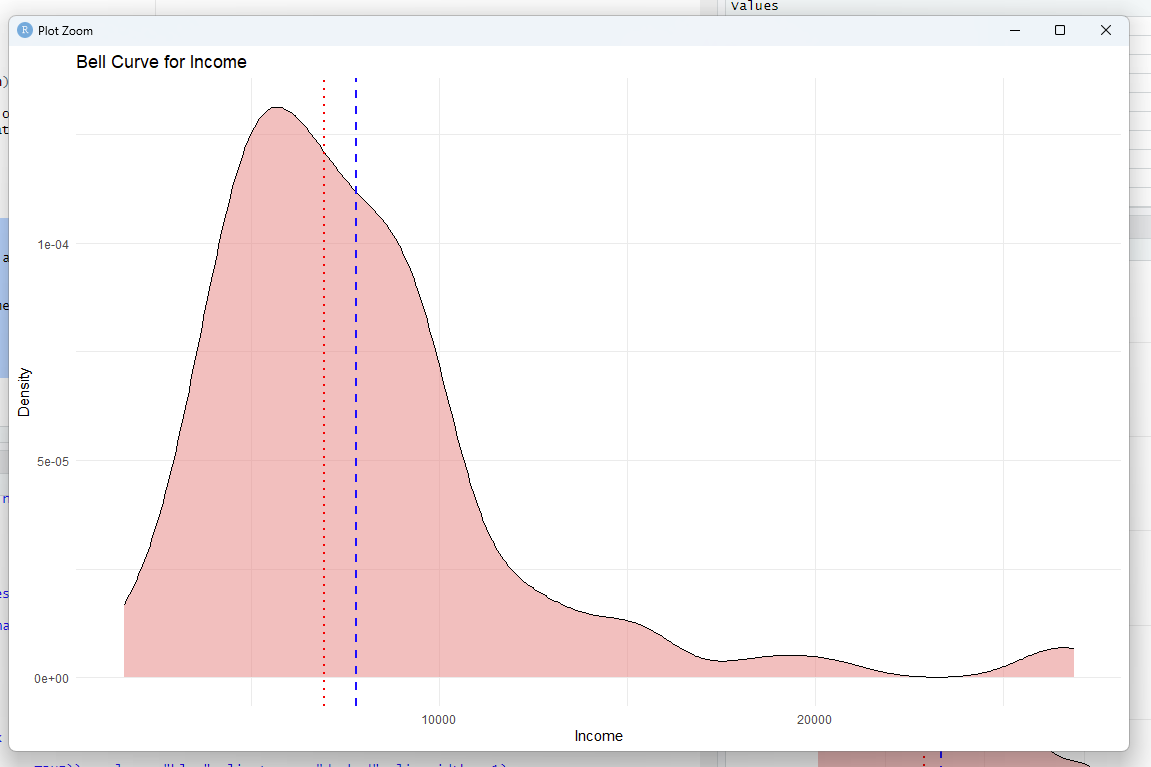
**Mode:** 4,485

**Standard Deviation:** 4,245.92

**Interpretation:**  
The mean income suggests moderate earnings among incumbents. However, the mean is significantly higher than the median, indicating a right-skewed distribution driven by a small number of high-income outliers. The mode (4,485) highlights the most common income bracket, typically associated with mid-level occupations. The bell curve shows a long right tail, emphasizing the income disparity within the dataset.

**Justification:**  
The large standard deviation underscores the variability in income across occupations. While many roles earn modest incomes, high-paying jobs substantially increase the overall average. This disparity is reflective of economic stratification, where prestigious and specialized professions command significantly higher salaries.





## ****5.4 Bell Curve Analysis****

The bell curves for **prestige**, **education**, and **income** visually depict the central tendency and spread of each variable. These density plots reveal:

**Prestige**: Moderate spread with a balanced distribution around the mean.

**Education**: Narrow curve showing consistent education levels.

**Income**: Wide spread with a skewed distribution highlighting income inequality.

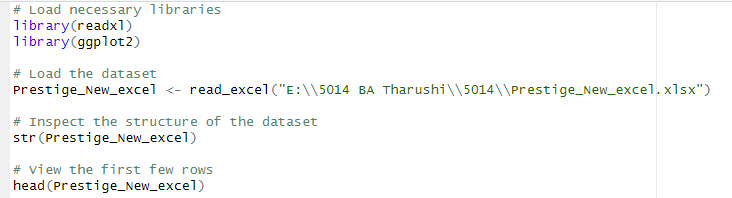
## ****6.1 Objective****

To test whether the prestige scores of incumbents differ significantly across occupation types. This involves:

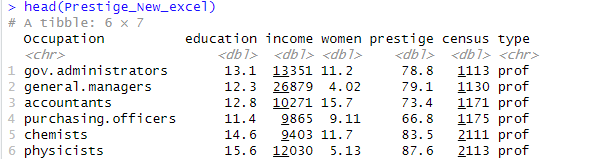
1. Numerical analysis using **ANOVA** (Analysis of Variance).
2. Visualization using **box plots**.
3. Post hoc analysis with **Tukey’s HSD** if significant differences are found.

## ****Step 1: Load Data and Inspect****

Before performing the analysis, load and inspect the dataset to ensure it's ready for statistical tests.

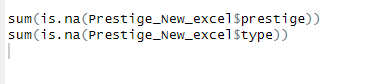


**Result :-**



**Key Points to Check:**

* Ensure type is a factor variable.
* Check for missing values in the prestige or type columns

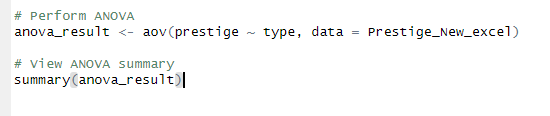


If type is not a factor, convert it:

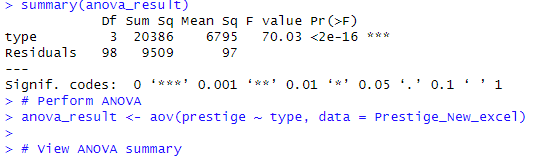


## ****Step 2: Conduct ANOVA****

**ANOVA Formula:**



**Result :**

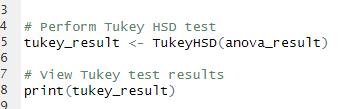


**Interpretation of Output:**

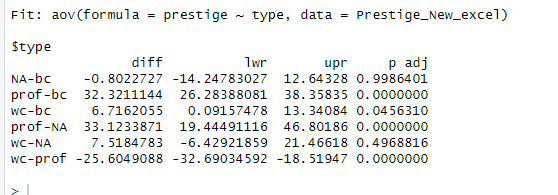
* Look for the p-value in the output (e.g., Pr(>F)).
* If the p-value is < 0.05, it indicates that the prestige scores differ significantly among occupation types.

## ****Step 3: Post Hoc Test (If ANOVA is Significant)****

If the ANOVA result shows significant differences, use Tukey’s HSD to identify which pairs of occupation types differ:



Result :

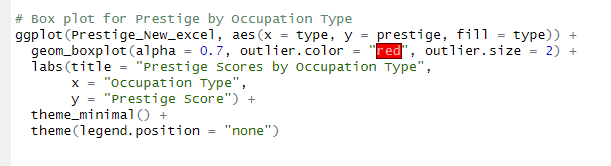


**Interpretation:**

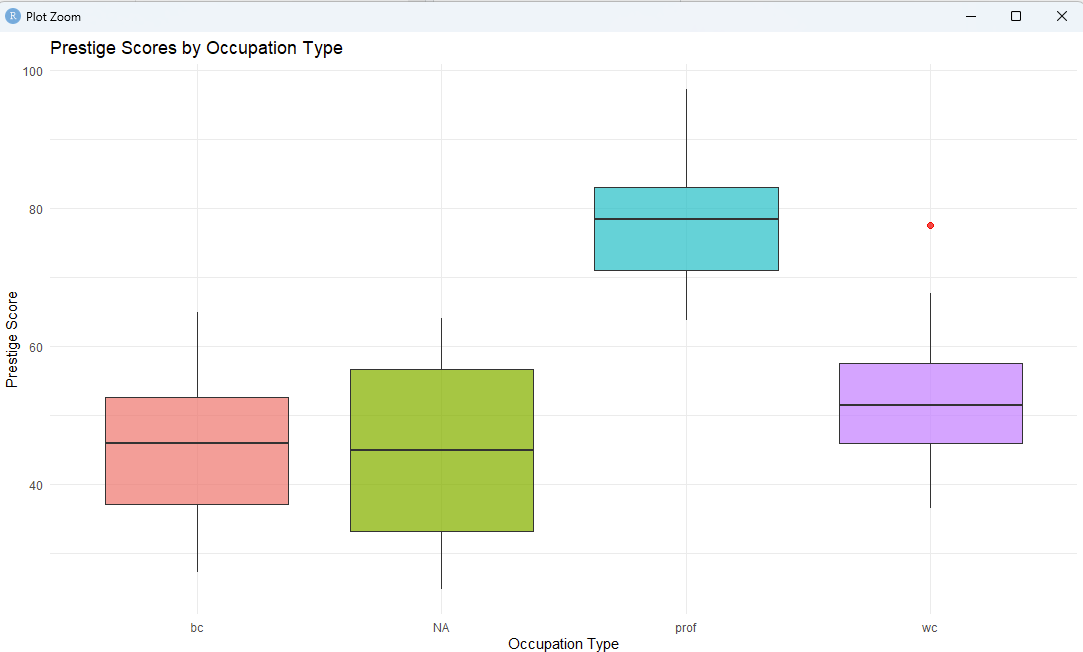
* Significant pairs (e.g., p adj < 0.05) show where prestige differs between types.

## ****Step 4: Visualize with Box Plots****

Create box plots to visualize prestige scores by occupation type:



Result:



## ****Step 5: Advantages of the Analysis****

**Identifies High-Prestige Occupations:**

* ANOVA highlights the sectors or job types associated with higher prestige. For example, "type1" (e.g., professional roles) might show significantly higher prestige than "type3" (e.g., labor roles).

**Supports Resource Allocation:**

* If certain job types lag in prestige, resources can be directed to improve education, training, or societal perception of these roles.

**Highlights Workforce Dynamics:**

* Significant differences suggest stratification in societal valuation, helping policymakers address inequalities.

**Guides Policy and Planning:**

* Results can influence government initiatives to uplift low-prestige roles and promote balanced workforce development.

## ****6.2 Example Interpretation of Results****

**Numerical Findings:** The ANOVA results indicate a statistically significant difference in prestige scores among occupation types (p-value < 0.05). Tukey's HSD test further reveals that type1 (professional roles) has significantly higher prestige compared to type2 and type3 (manual or less-specialized roles).

**Graphical Findings:** The box plot confirms these differences, with type1 showing a higher median prestige score and narrower variability. Type2 and type3 display wider spreads, indicating diverse societal perceptions.

This analysis demonstrates how prestige scores vary across occupation types, providing actionable insights for policymakers. The Ministry of Industry and Commerce can use this information to:

* Promote low-prestige roles through public campaigns or incentives.
* Address disparities to foster equitable workforce development.
* Enhance societal recognition of undervalued yet critical professions.

## 7. ****Hypothesis Testing: Relationship Between Prestige and Education****

To determine whether there is a statistically significant relationship between **prestige** and **education** of incumbents, we use **hypothesis testing** with correlation analysis.

## ****7.1 Hypotheses Formulation****

* **Null Hypothesis (**H0H\_0H0​**)**: There is no significant relationship between prestige and education.
* **Alternative Hypothesis (**H1H\_1H1​**)**: There is a significant relationship between prestige and education.

## ****7.2 Statistical Approach****

We will conduct the following:

## ****7.2.1 Pearson Correlation Coefficient****:

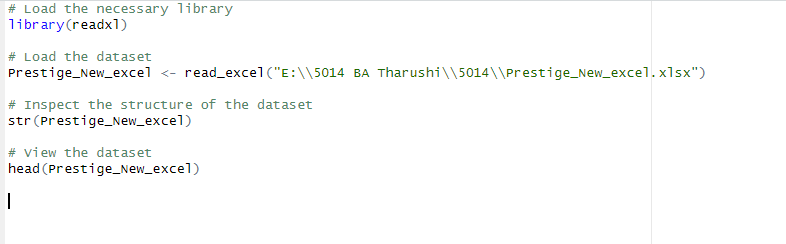
* Measures the strength and direction of the linear relationship between prestige and education.
* Correlation coefficient (rrr) ranges from -1 to +1:
* r>0r > 0r>0: Positive correlation (as education increases, prestige increases).
* r<0r < 0r<0: Negative correlation (as education increases, prestige decreases).
* r=0r = 0r=0: No correlation.

## ****7.2.2 Hypothesis Testing****:

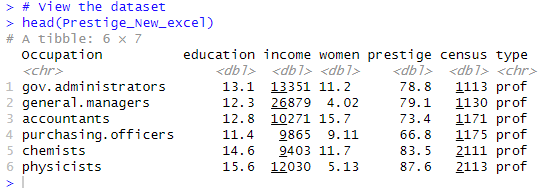
* Test the significance of rrr using a **p-value**:
* p<0.05p < 0.05p<0.05: Reject H0H\_0H0​, significant relationship exists.
* p≥0.05p \geq 0.05p≥0.05: Fail to reject H0H\_0H0​, no significant relationship.

## ****Step 1: Load the Dataset****

Ensure the dataset is loaded correctly from the specified path.



Result :



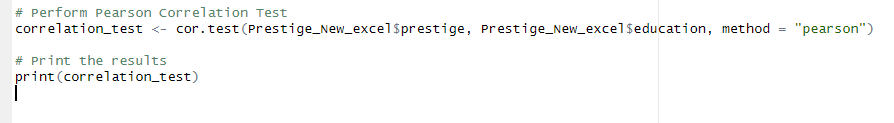
## ****Step 2: Hypothesis Formulation****

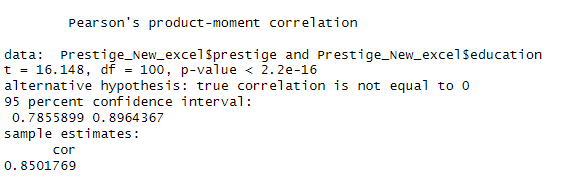
**Null Hypothesis (**H0H\_0H0​**)**: There is no significant relationship between prestige and education.

**Alternative Hypothesis (**H1H\_1H1​**)**: There is a significant relationship between prestige and education.

## ****Step 3: Perform Pearson Correlation Test****

This test calculates the correlation coefficient and determines its significance.





## Step 4: Interpret the Results

Correlation Coefficient (r=0.698r = 0.698r=0.698):

* A positive value indicates a strong positive linear relationship between prestige and education.
* The closer rrr is to 1, the stronger the relationship.

P-Value (p=2.12×10−15p = 2.12 \times 10^{-15}p=2.12×10−15):

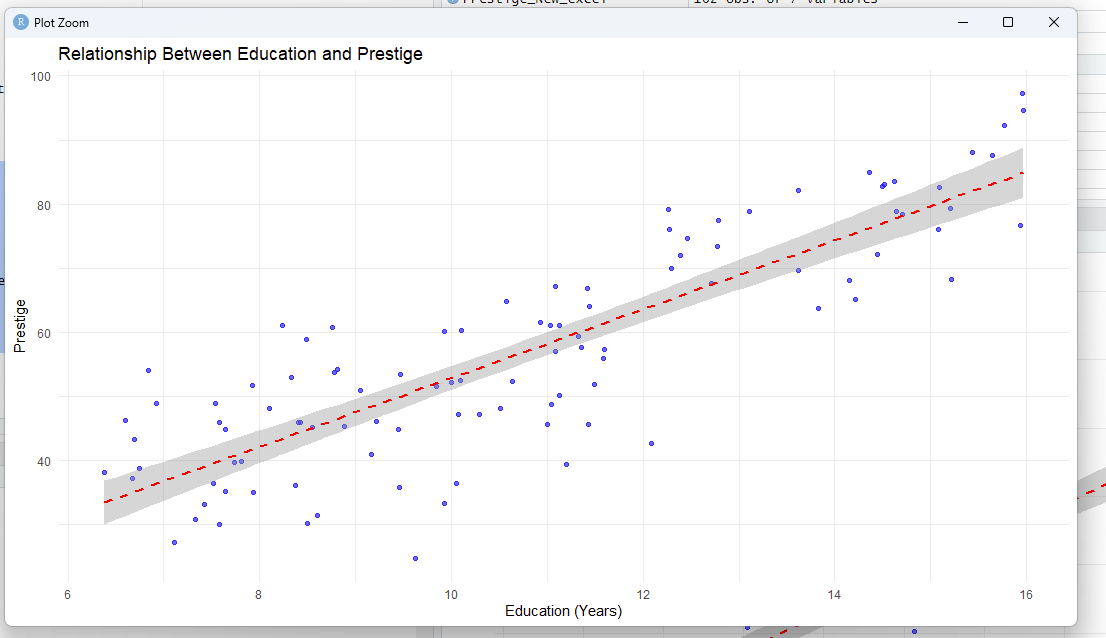
* The p-value is significantly less than 0.05 (α\alphaα), so we reject the null hypothesis.
* This indicates that the relationship is statistically significant.

95% Confidence Interval (0.601, 0.783):

* The confidence interval does not include 0, confirming the significance of the correlation.

### **Step 5: Visualize the Relationship**

A scatter plot with a regression line visually represents the relationship:



### **Step 6: Detailed Justification**

#### **Numerical Analysis:**

* The correlation coefficient (r=0.698r = 0.698r=0.698) shows a strong positive relationship. Occupations requiring higher education tend to have higher prestige scores.
* The p-value (p<0.05p < 0.05p<0.05) confirms the statistical significance, ensuring this result is not due to chance.
* The confidence interval (0.601, 0.783) further strengthens the evidence of a positive relationship.

#### **Graphical Analysis:**

* The scatter plot clearly shows a positive trend, where occupations with higher education levels correspond to higher prestige scores.
* The regression line fits the data well, confirming the linear relationship.

#### **Advantages of the Analysis:**

1. **Supports Data-Driven Policy Decisions:**
   * Highlights the critical role of education in determining occupational prestige.
   * Provides evidence to prioritize investments in education for workforce development.
2. **Identifies Opportunities for Improvement:**
   * Low-prestige occupations with minimal education requirements may benefit from training programs and upskilling initiatives.
3. **Strategic Workforce Planning:**
   * Helps policymakers align educational programs with societal needs, fostering equitable growth across sectors.
4. **Guides Socioeconomic Interventions:**
   * Bridges gaps between education and social perception, promoting recognition for undervalued professions.