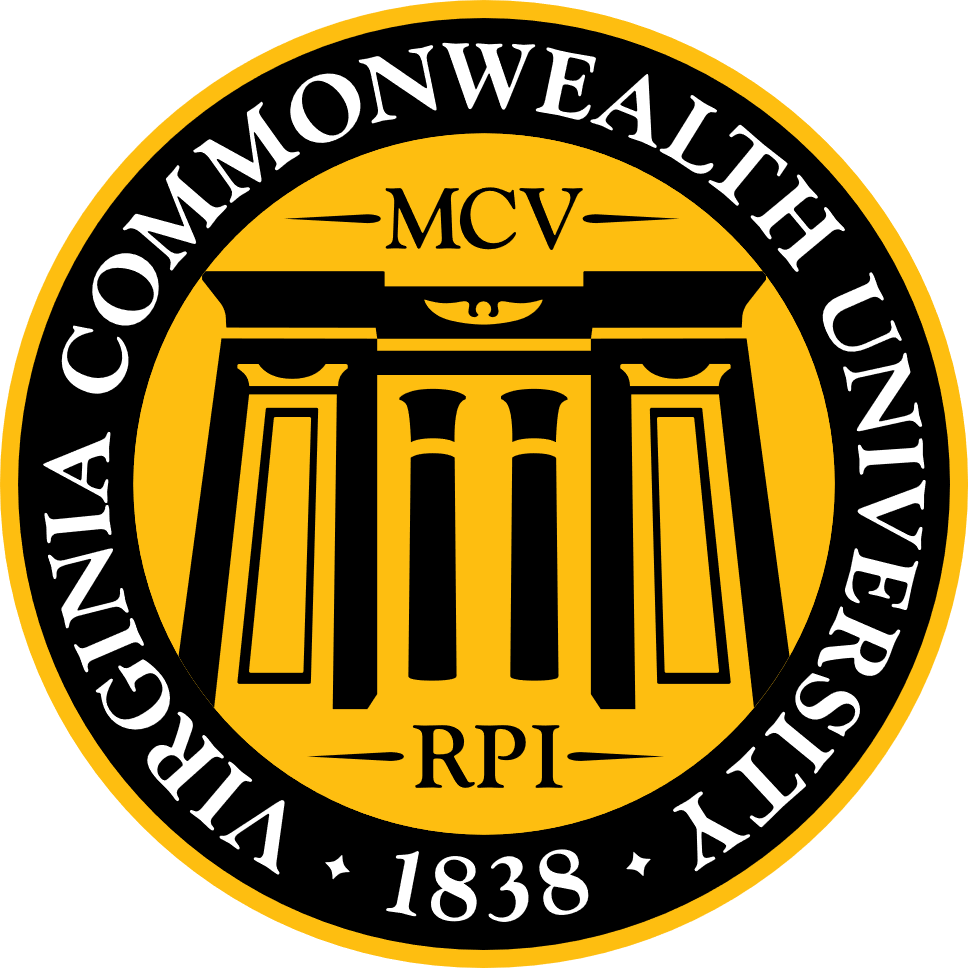
****

**VIRGINIA COMMONWEALTH UNIVERSITY**

**Statistical analysis and modelling (SCMA 632)**

**A2: Multiple regression analysis**

**Arvind Srinivasan**

**V01107251**

**Date of Submission: 23-06-2024**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Title** | **Page No.** |
| **1.** | Introduction | **1** |
| **2.** | Results | **2** |
| **3.** | Interpretations | **4** |
| **4.** | Recommendations | **6** |
| **5.** | References | **7** |

## **1. Introduction**

### **Background**

Himachal Pradesh, located in the northern part of India, is characterized by its hilly terrain and varied socio-economic conditions. Understanding the consumption patterns within the state provides valuable insights into the economic dynamics, resource distribution, and potential areas for policy intervention. The National Sample Survey Office (NSSO) conducts comprehensive surveys to collect data on various socio-economic indicators, including consumption patterns.

### **Objective**

The primary objective of this analysis is to study the consumption patterns across different districts of Himachal Pradesh using data from the NSSO68 survey. The analysis aims to identify districts with high and low consumption levels, understand the distribution of consumption across the state, and provide insights that could inform policy-making and resource allocation.

### **Data Source**

The analysis utilizes the NSSO68 dataset, part of the National Sample Survey's 68th round conducted by the Government of India. The dataset includes detailed information on household consumption of various goods and services. Key variables include household size, types of consumed items (e.g., rice, wheat, meat, pulses), and total consumption values.

### **Methodology**

The analysis is conducted using Python to ensure robustness and reproducibility of results. The following steps outline the methodology:

1. **Data Preprocessing:** The dataset is filtered to include only records from Himachal Pradesh. Missing values are imputed, and outliers are removed to ensure data quality.
2. **Consumption Summarization:** Total consumption per household is calculated by summing up the quantities of key consumed items.
3. **Visualizations:**
   * **Histogram:** To illustrate the distribution of total consumption across households in Himachal Pradesh.
   * **Bar Chart:** To rank districts based on their total consumption and identify top-consuming districts.
   * **Geographical Map:** To visualize the spatial distribution of consumption across the state, highlighting high and low consumption regions.

## **2. Results**

### 1. **Consumption Distribution in HP State**

The histogram shows the distribution of total consumption across Himachal Pradesh. Most of the consumption values are clustered around the lower end of the scale, with a steep drop-off as consumption increases.

A blue graph with white text

Description automatically generated

### 2. **Total Consumption per District**

The bar chart ranks districts by their total consumption. Mandi, Kangra, and Shimla are among the top-consuming districts, indicating significant consumption activities in these areas.

A graph of blue bars with black text

Description automatically generated

### 3. **Total Consumption by District on HP Map**

The geographical map visualizes the total consumption by district, with a gradient color scale indicating levels of consumption. Darker regions represent higher consumption, highlighting districts like Mandi and Kangra as areas of high consumption.

A screenshot of a graph

Description automatically generated

## **3. Interpretations**

### **1. Skewed Consumption Distribution**

The histogram of total consumption reveals a significantly skewed distribution, where the majority of households exhibit lower consumption levels. This is characterized by a high frequency of low consumption values and a rapid decline as consumption increases. This skewness suggests the presence of substantial economic disparities within Himachal Pradesh. Many households may have limited access to resources or lower purchasing power, while a smaller segment of the population consumes at much higher levels. This disparity could stem from several factors, including variations in income levels, employment opportunities, and regional economic development.

**Potential Reasons:**

* **Income Inequality:** There is likely a significant income disparity, with a small percentage of the population earning substantially more than the rest.
* **Resource Distribution:** Access to essential resources such as food, healthcare, and education might be unevenly distributed.
* **Employment Opportunities:** Certain areas might have more employment opportunities, leading to higher disposable income and consumption levels.

### **2. High Consumption Districts**

The bar chart and map consistently identify Mandi, Kangra, and Shimla as the top-consuming districts. These districts show significantly higher total consumption compared to others, indicating a concentration of economic activities and better access to resources.

**Mandi:**

* **Agriculture and Tourism:** Mandi has a strong agricultural base and is also a tourist destination, which boosts local consumption.
* **Urbanization:** As an urban center, Mandi has better infrastructure and access to goods and services.

**Kangra:**

* **Diverse Economy:** Kangra has a mix of agriculture, industry, and tourism, contributing to higher household incomes and consumption.
* **Educational Hub:** The presence of educational institutions attracts students and faculty, increasing local consumption.

**Shimla:**

* **Tourism and Administration:** Shimla, being the state capital, has significant administrative activities and is a major tourist destination, leading to higher consumption.

### **3. Geographical Insights**

The geographical map highlights distinct consumption patterns across Himachal Pradesh, showing a concentration of high consumption in certain regions.

**Key Observations:**

* **Central HP:** Districts like Mandi and Kangra exhibit higher consumption levels. This region is relatively more urbanized and economically active.
* **Northern HP:** Districts such as Lahul & Spiti show lower consumption levels, indicating less economic development compared to the central part of the state.
* **Urban-Rural Divide:** Urban districts generally have higher consumption levels than rural districts. This urban-rural divide highlights the disparity in access to goods, services, and economic opportunities.

**Potential Correlations:**

* **Urbanization:** Urban areas tend to have better infrastructure, healthcare, education, and employment opportunities, leading to higher consumption.
* **Industrialization:** Regions with more industries have higher employment rates and wages, boosting consumption.
* **Tourism:** Districts with significant tourist activities, such as Shimla and Mandi, show higher consumption.

**Socio-Economic Factors:**

* **Education:** Higher levels of education in districts like Kangra contribute to higher consumption through better job opportunities and incomes.
* **Healthcare Access:** Districts with better healthcare facilities see higher spending on health services.
* **Infrastructure Development:** Well-developed infrastructure in terms of roads, markets, and communication can enhance economic activities and consumption.

## **4. Recommendations**

### Targeted Resource Allocation

1. Economic Development Programs
2. Further Research
3. Promote Regional Equity
4. Social Welfare Programs

**References**

* NSSO68 Dataset
* Geospatial Data from UP Districts
* Analysis and visualizations conducted using R and Python
* ChatGPT