This document provides detailed step-by-step instructions for conducting data preparation, Principal Component Analysis (PCA) in SPSS, and Geographically Weighted Regression (GWR) in ArcGIS Pro.

**Data source:** We obtained LCM cases from CDC WONDER Wide-ranging Online Data<https://wonder.cdc.gov/controller/datarequest/D77>, Environmental factors, including PM2.5, Temperature, Wind, NO2, Ozone, SO2, CO were from U.S. Environmental Protection Agency, Smoking rate derived from U.S. Centers for Disease Control and Prevention. The rank of GDP was from <https://www.bea.gov/data/gdp>. Health conditions, such as public health, public health access, and Health care quality, were collected by <https://www.usnews.com/> .

**Data procedure:**

1. **Go google drive to open google colab and upload raw data.** All raw data on county level should be save one excel spreadsheet
2. Identify the zone numbers, run the code, and save the results.
3. partition all variable on your selected zone numbers and make sure each variables were split on your selected zone numbers.
4. Standardization all data on the 455, 905, 1306 (county) levels.
5. Run Principal Component Analysis in SPSS to avoid mulicolinearity.

**Step 1: Prepare Your Data**

1. **Input Data**:
   * Ensure your dataset is loaded into SPSS.
   * Variables should be numeric and measured on comparable scales. If necessary, standardize variables (e.g., z-scores) before analysis.
2. **Check for Missing Values**:
   * Handle missing data by imputation or listwise/pairwise deletion to ensure consistency in analysis.
3. **Suitability for PCA**:
   * Conduct preliminary checks to confirm that PCA is appropriate:
     + **Kaiser-Meyer-Olkin (KMO) Test**: Measures sampling adequacy (value > 0.6 is recommended).
     + **Bartlett’s Test of Sphericity**: Tests whether the correlation matrix is suitable for PCA (significance < 0.05 indicates suitability).

**Step 2: Access the PCA Procedure**

1. Go to the **Analyze** menu in SPSS.
2. Select **Dimension Reduction**, then click **Factor**.

**Step 3: Configure PCA Settings**

1. **Input Variables**:
   * In the dialog box, move all the variables you want to include in the PCA into the **Variables** box.
2. **Extraction Method**:
   * Click the **Extraction** tab.
   * Under **Method**, select **Principal Components**.
   * Choose either:
     + **Correlation Matrix** (default): For variables with different units/scales.
     + **Covariance Matrix**: For variables measured on the same scale.
   * Specify the number of components to extract, or let SPSS determine them automatically using:
     + **Eigenvalues > 1** (default).
     + A fixed number of components (if theory dictates).
3. **Rotation** (Optional):
   * Click the **Rotation** tab.
   * Apply a rotation method (e.g., **Varimax** for orthogonal rotation or **Promax** for oblique rotation) to simplify interpretation.
4. **Scores**:
   * To save component scores, click the **Scores** tab and select **Save as Variables**.
5. **Statistics**:
   * Click the **Statistics** tab.
   * Check **KMO and Bartlett’s test** and **Anti-image** to evaluate sampling adequacy and correlation structure.
   * Select **Reproduced Correlations** if needed.
6. **Plots** (Optional):
   * Click the **Plots** tab and select **Scree Plot** to visualize eigenvalues and determine the number of components.

**Step 4: Run PCA**

* Click **OK** to run the PCA.

**Step 5: Interpret the Results**

1. **Output Window**:
   * **KMO and Bartlett’s Test**: Confirm data suitability for PCA.
   * **Total Variance Explained**:
     + Check eigenvalues to identify significant components.
     + Look at the cumulative percentage of variance explained to decide how many components to retain.
   * **Scree Plot**: Identify the "elbow" point to determine the number of components visually.
   * **Component Matrix**: Examine the loadings of variables on components to interpret their relationships.
   * **Rotated Component Matrix**: Interpret simpler, more meaningful patterns if rotation was applied.
2. **Saved Scores**:
   * PCA scores (if saved) will appear as new variables in the dataset, representing the principal components.
3. **Run Geographical Weighted Regression** into ArcGIS pro
4. **Input Data**:
   * Ensure your dataset contains both dependent (response) and independent (explanatory) variables in a spatial format (e.g., shapefile, geodatabase feature class).
   * Include geographic coordinates for spatial locations.
5. **Projection**:
   * Project your data into a coordinate system that uses linear units (e.g., meters), as GWR calculations depend on distances.
6. **Check Data Requirements**:
   * No missing values in the dependent or explanatory variables.
   * A numeric field to represent your dependent variable.
7. **Access the Geoprocessing Tool**
8. Open **ArcGIS Pro** and load your data.
9. Navigate to the **Analysis** tab and click **Tools** to open the Geoprocessing pane.
10. Search for the **Geographically Weighted Regression (GWR)** tool.
11. **Configure the GWR Tool**
12. **Input Feature Class**:
    * Select the feature layer containing your data.
13. **Dependent Variable**:
    * Choose the variable you are trying to predict (response variable).
14. **Explanatory Variables**:
    * Select one or more variables that might influence the dependent variable.
15. **Kernel Type**:
    * Select a kernel type:
      + **Fixed**: Uses a constant bandwidth distance.
      + **Adaptive**: Adjusts the bandwidth based on local feature density.
16. **Bandwidth Method**:
    * Choose a method for bandwidth selection:
      + **AICc** (default): Optimizes model performance.
      + **User-defined**: Specify your own bandwidth.
17. **Output Fields and Residuals**:
    * Choose to output fields like coefficients, standard errors, and residuals.
18. **Output Feature Class**:
    * Specify the location to save the output feature class with GWR results.
19. (Optional) **Diagnostic Information**:
    * Select the checkbox for diagnostics like R², AICc, and residual sum of squares.
20. **Run the Tool**
21. Click **Run** to execute the GWR analysis.
22. The tool will output:
    * A feature layer with GWR coefficients and diagnostics for each feature.
    * Optional diagnostic messages in the Geoprocessing pane.
23. **Interpret the Results**
24. Open the output feature class attribute table.
25. Examine the coefficients for each explanatory variable and the diagnostic statistics.
    1. **Local coefficients**: Represent the strength and direction of relationships in different spatial areas.
    2. **Residuals**: Evaluate model fit.
    3. **AICc/R²**: Compare with other models or zone configurations to assess performance.
26. **Visualize Results**
27. Create thematic maps of the local coefficients to visualize spatial patterns.
28. Use the **Symbology** pane to display residuals and identify areas where the model fits poorly.

Please follow the instructions provided above to replicate the analysis and achieve results consistent with mine. Thank you.