|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Assumptions (Pearson and Predictive Correlation)** | **How assumption is tested/determined** | **If test statistic - Value** | **Significance**  **(p-value)** | **Assumption met?**  **Yes or No** |
| Linear Relationship | Scatterplot of interval or ratio data |  |  | Yes |
| Outliers assessment | Exploratory statistics (PP and QQ plot) Standardize and values > 3.29 = outliers |  |  | Yes |
| Independence of residuals (no autocorrelation) | Durbin-Watson’s d test. Null = no autocorrelation. | 1.765 | Critical value range 1.76-1.78 (n=200)  OR  1.5 < d < 2.5 | Yes (accept null) |
| Homoscedasticity | Goldfeld-Quandt Test (python) or Visual inspection of scatterplot |  |  | Yes |
| Residuals are normally distributed | Histogram, P-P or Q-Q plot |  |  | Yes |

# Linear Regression (Predictive models) Assumption Flow

Step 1

1. Linear relationships between variables
2. No significant outliers

Step 2

1. Step 1 fail:
   * Corrections (remove outliers)
   * Re-assessment of step 1
2. Step 1 Pass:
   * Move to step 3

Step 3

1. Check for independence of errors (residuals) or autocorrelation
2. Homoscedasticity of residuals (variance of residuals should be constant)
3. Residuals are normally distributed

# **Linear Regression Flow Chart**

1. Linear relationship of variables
2. No significant or extreme outliers
3. Check for independence of errors (residuals) or autocorrelation
4. Homoscedasticity of residuals
5. Residuals are normally distributed

Fail

Pass

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Assumptions (Spearman Rank)** | **How assumption is tested/determined** | **If test statistic - Value** | **Significance**  **(p-value)** | **Assumption met?**  **Yes or No** |
| Continuous or ordinal variables | Inherent in data |  |  | Yes |
| Paired variables per subject | Inherent in data |  |  | Yes |
| Monotonic relationship between the 2 variables | Scatterplot showing that data only takes one slope direction |  | Sloped and possibly not linear (outliers accepted) | Yes |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Assumptions (Spearman Rank)** | **How assumption is tested/determined** | **If test statistic - Value** | **Significance**  **(p-value)** | **Assumption met?**  **Yes or No** |
| 1 continuous vs 1 dichotomous | Inherent in data |  |  | Yes |
| Normal continuous variable | Histogram and P-P plot |  |  | Yes |
| Homogeneity of variance met | Lavene’s test |  |  | Yes |