



Parkinson's Disease Dataset Analysis

The purpose of this research paper is to provide a comprehensive, step-by-step guide for conducting **descriptive statistics** using the Isalos Analytics Platform. Descriptive statistics summarise the main characteristics of the patient cohort.

These methods are fundamental in healthcare research for identifying demographic and lifestyle factors associated with disease severity, cognitive performance, and functional impairment.

Isalos version used: 2.1.5

The dataset used in this analysis is the **Parkinson's Disease Dataset Analysis**, available on Kaggle.

Source: [Parkinson's Disease Dataset Analysis](#)

This dataset simulates real-world healthcare records and includes comprehensive information for **2,105** patients diagnosed with Parkinson's Disease. It contains demographic details (age, gender, ethnicity, education), lifestyle factors (smoking, alcohol consumption, physical activity, diet quality), and clinical measurements (UPDRS, MoCA, Functional Assessment). It is designed to support predictive modelling and risk assessment for Parkinson's disease progression and cognitive decline.

Step 1: Import Data from file

Firstly, launch the Isalos Analytics Platform (version 2.1.5). Then, in the left spreadsheet panel, create a new tab by clicking **File** and choose **Open** to load the Parkinson's Disease dataset (or simply copy-paste the data from your spreadsheet).

Col1	Col2 (I)	Col3 (I)	Col4 (I)	Col5 (I)	Col6 (D)	Col7 (I)	Col8 (D)	Col9 (D)	Col10 (D)	
User Row ID	Age	Gender	Ethnicity	EducationLevel	BMI	Smoking	AlcoholConsumption	PhysicalActivity	DietQuality	
3058	85	0	3	1	19.61987796	0	5.108240607	1.380659917	3.893969135	
3059	75	0	0	2	16.24733916	1	6.027648029	8.40980405	8.51342825	
3060	70	1	0	0	15.36823871	0	2.242135331	0.213274591	6.498804606	
3061	52	0	0	0	15.45455733	0	5.997787563	1.375045164	6.715033333	
3062	87	0	0	1	18.61604177	0	9.775242923	1.188607062	4.657572037	
3063	68	1	2	1	39.42331141	1	13.5968889	7.796704004	7.070238878	
3064	78	1	0	0	30.54200329	1	2.011281313	9.028536304	9.838445926	
3065	70	1	0	0	36.75828161	1	19.98886597	3.891748622	3.421960006	
3066	80	0	2	1	22.3805865	1	7.293287715	2.595670177	4.784827139	
3067	71	0	3	2	23.72708628	1	17.78290985	7.344890316	3.393018462	
3068	70	0	0	3	38.48254474	0	6.639761993	7.872186697	9.225710184	
3069	53	0	0	1	35.8967387	1	5.212906263	7.185203018	7.918912222	
3070	74	0	2	2	30.22551208	0	3.762918066	4.316651243	5.112520425	
3071	87	1	1	2	38.29830655	0	12.61599522	9.299289996	6.71557911	
3072	58	0	0	3	34.96532287	1	11.70859735	4.392463152	5.182038871	
3073	56	1	0	0	18.95878125	1	2.04712047	9.432830343	1.720277262	
3074	54	1	0	0	28.04958007	1	7.260733832	4.311617423	2.509936233	

Step 2: Access the Descriptive Statistics and ANOVA functions

Navigate to the descriptive statistics module through the main menu by choosing:

Statistics → Basic Statistics → Descriptive Statistics

Then you will see the descriptive statistics configuration window.


Descriptive Statistics
?
×

Select Type
Sample

1. Sample Size & Counts

☐ Size/Count (n)
☐ Frequency
☐ Sum (Σ)
☐ Sum of Squares (SS)

2. Central Tendency

☐ Mean ($\bar{x}$)
☐ Median ($\tilde{x}$)
☐ Mode
☐ Midrange (MR)

3. Dispersion / Spread

☐ Standard Deviation (s)
☐ Variance (s^2)
☐ Range (R)
☐ Mean Absolute Deviation (MAD)
☐ Minimum (min)
☐ Maximum (max)

4. Quartiles, UQR & Outliers

Method: ☐ Inclusive ☐ Exclusive
☐ Quartiles (Q1, Q2, Q3)
☐ Interquartile Range (IQR)
☐ Outliers

5. Relative / Advanced Measures

☐ Coefficient of Variation (CV)
☐ Relative Standard Deviation (RSD)
☐ Root Mean Square (RMS)
☐ Standard Error of the Mean (SE $\bar{x}$)

6. Distribution Shape

☐ Skewness (γ_1)
☐ Kurtosis (β_2)
☐ Kurtosis Excess (α_2)

7. Normality Tests

☐ Shapiro-Wilk
☐ Kolmogorov-Smirnov
☐ Anderson-Darling

8. Plots

☐ Q-Q Plot
☐ Histogram
☐ Box Plot
☐ Violin Plot

Execute
Cancel

Step 3: Configure parameters

In the descriptive statistics window, set the following parameters:

1. Sample Size & Counts
 - a. Size / Count (n)
 - b. Frequency
 - c. Sum
 - d. Sum of Squares
2. Central Tendency
 - a. Mean
 - b. Median
 - c. Mode
 - d. Midrange
3. Dispersion / Spread
 - a. Standard Deviation
 - b. Variance
 - c. Range
 - d. Mean Absolute Deviation
 - e. Minimum

- f. Maximum
- 4. Quartiles, UQE & Outliers
 - a. Quartiles (Q1,Q2,Q3)
 - b. Interquartile Range (IQR)
 - c. Outliers
- 5. Relative / Advanced Measures
 - a. Coefficient of Variation
 - b. Relative Standard Deviation
 - c. Root Mean Square
 - d. Standard Error of the Mean
- 6. Distribution Shape
 - a. Skewness
 - b. Kurtosis
 - c. Kurtosis excess
- 7. Normality Tests
 - a. Shapiro-Wilk
 - b. Kolmogorov-Smirnov
 - c. Anderson-Darling
- 8. Plots
 - a. Q-Q Plot
 - b. Histogram
 - c. Box Plot
 - d. Violin Plot

The above preview will display the multivariate model as shown below:

Descriptive Statistics

Select Type

Sample

1. Sample Size & Counts

☒ Size/Count (n)
 ☒ Frequency
 ☒ Sum (Σ)
 ☒ Sum of Squares (SS)

2. Central Tendency

☒ Mean ($\bar{x}$)
 ☒ Median ($\tilde{x}$)
 ☒ Mode
 ☒ Midrange (MR)

3. Dispersion / Spread

☒ Standard Deviation (s)
 ☒ Variance (s^2)
 ☒ Range (R)
 ☒ Mean Absolute Deviation (MAD)
 ☒ Minimum (min)
 ☒ Maximum (max)

4. Quartiles, UQR & Outliers

Method: ☒ Inclusive ☒ Exclusive
 ☒ Quartiles (Q1, Q2, Q3)
 ☒ Interquartile Range (IQR)
 ☒ Outliers

5. Relative / Advanced Measures

☒ Coefficient of Variation (CV)
 ☒ Relative Standard Deviation (RSD)
 ☒ Root Mean Square (RMS)
 ☒ Standard Error of the Mean (SE $\bar{x}$)

6. Distribution Shape

☒ Skewness (γ_1)
 ☒ Kurtosis (β_2)
 ☒ Kurtosis Excess (α_2)

7. Normality Tests

☒ Shapiro-Wilk
 ☒ Kolmogorov-Smirnov
 ☒ Anderson-Darling

8. Plots

☒ Q-Q Plot
 ☒ Histogram
 ☒ Box Plot
 ☒ Violin Plot

Distribution:

Normal

Execute

Cancel

After configuring all parameters:

- * Review all selections to ensure accuracy.
- * Click the **Execute** button to run the descriptive analysis.
- * Wait for the computation to complete (processing time depends on dataset size and the number of dependent variables).

Step 4: Results

The analysis results will be displayed in the right spreadsheet panel. The output typically includes the following components:

Factors	Size/Count	Sum	Sum of Squares	Mean	Median	Mode	Mid Range	Standard Deviation	Variance
Age	2105.0	146512.0	282846.3923990	69.6019002	70.0	84.0	69.5	11.5945115	134.4326960

Range	Mean Absolute Deviation	Minimum	Maximum	Quartiles - Inclusive	Quartiles - Exclusive	Interquartile Range - Inclusive	Interquartile Range - Exclusive		
39.0	9.9904492	50.0	89.0	Q1: 60.000000, Q2: 70.000000, Q3: 80.000000	Q1: 60.000000, Q2: 70.000000, Q3: 80.000000	20.0	20.0		
Coefficient Of Variation	Relative Standard Deviation	Root Mean Square	Standard Error Of Mean	Skewness	Kurtosis	Kurtosis Excess	Shapiro-Wilk: W	Shapiro-Wilk: p-value	Shapiro-Wilk: Prob
0.1665833	16.6583260	70.560565	0.2527123	-0.0334598	1.8160222	-1.1839778	0.9534394	0.0	1.0
Shapiro-Wilk: p-value	Shapiro-Wilk: Prob	Shapiro-Wilk: Decision	Anderson-Darling: A ²	Anderson-Darling: p-value	Anderson-Darling: Prob	Anderson-Darling: Decision	Kolmogorov-Smirnov: D	Kolmogorov-Smirnov: p-value	Kolmogorov-Smirnov: Prob
0.0	1.0	Reject Null Hypothesis	0.0691501	0.0	1.0	Reject Null Hypothesis	23.2069780	0E-7	1.0
Shapiro-Wilk: Decision	Anderson-Darling: A ²	Anderson-Darling: p-value	Anderson-Darling: Prob	Anderson-Darling: Decision	Kolmogorov-Smirnov: D	Kolmogorov-Smirnov: p-value	Kolmogorov-Smirnov: Prob	Kolmogorov-Smirnov: Decision	
Reject Null Hypothesis	0.0691501	0.0	1.0	Reject Null Hypothesis	23.2069780	0E-7	1.0	Reject Null Hypothesis	





