



Heart Failure Prediction

The purpose of this research paper is to provide a comprehensive, step-by-step guide for conducting a complete statistical analysis workflow using the Isalos Analytics Platform. The process begins with **descriptive statistics**, which are used to summarize the main characteristics of the patient cohort, including demographic information and clinical measurements.

Isalos version used: 2.1.5

The dataset used in this analysis is the **Heart Disease Dataset** (UCI), available on Kaggle.

Source: [Heart Disease UCI](#)

This dataset simulates real-world healthcare records and includes patient-level information for over 900 individuals. It contains demographic details (age, sex), clinical measurements (resting blood pressure, cholesterol, maximum heart rate), and lifestyle/symptom indicators (chest pain type, exercise angina). It is widely used for predictive modelling and risk assessment for cardiovascular disease.

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 heart dataset (or simply copy-paste the data from your spreadsheet).

Age	Sex	ChestPainType	RestingBP	Cholesterol	FastingBS	RestingECG	MaxHR	ExerciseAngina	Oldpeak	ST_Slope	HeartDisease
40	M	ATA	140	289	0	Normal	172	N	0	Up	0
49	F	NAP	160	180	0	Normal	156	N	1	Flat	1
37	M	ATA	130	283	0	ST	98	N	0	Up	0
48	F	ASY	138	214	0	Normal	108	Y	1.5	Flat	1
54	M	NAP	150	195	0	Normal	122	N	0	Up	0
39	M	NAP	120	339	0	Normal	170	N	0	Up	0
45	F	ATA	130	237	0	Normal	170	N	0	Up	0
54	M	ATA	110	208	0	Normal	142	N	0	Up	0
37	M	ASY	140	207	0	Normal	130	Y	1.5	Flat	1
48	F	ATA	120	284	0	Normal	120	N	0	Up	0
37	F	NAP	130	211	0	Normal	142	N	0	Up	0
58	M	ATA	136	164	0	ST	99	Y	2	Flat	1
39	M	ATA	120	204	0	Normal	145	N	0	Up	0
49	M	ASY	140	234	0	Normal	140	Y	1	Flat	1
42	F	NAP	115	211	0	ST	137	N	0	Up	0

Step 2: Access the Descriptive Statistics function

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

Grouped by

Sex

ChestPainType

RestingECG

ExerciseAngina

1. Sample Size & Counts

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

4. Quartiles, UQR & Outliers

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

7. Normality Tests

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

2. Central Tendency

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

5. Relative / Advanced Measures

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

8. Plots

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

3. Dispersion / Spread

☐ Standard Deviation (s)
☐ Variance (s²)
☐ Range (R)
☐ Mean Absolute Deviation (MAD)
☐ Minimum (min)
☐ Maximum (max)

6. Distribution Shape

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

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 preview below displays the multivariate model:

Descriptive Statistics

Select Type

Sample

Grouped by

Sex

ChestPainType

RestingECG

ExerciseAngina

1. Sample Size & Counts

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

4. Quartiles, UQR & Outliers

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

7. Normality Tests

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

2. Central Tendency

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

5. Relative / Advanced Measures

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

8. Plots

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

3. Dispersion / Spread

☒ Standard Deviation (s)
 ☒ Variance (s²)
 ☒ Range (R)
 ☒ Mean Absolute Deviation (MAD)
 ☒ Minimum (min)
 ☒ Maximum (max)

6. Distribution Shape

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

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: (some of them)

Sex	ChestPainType	RestingECG	ExerciseAngina	ST_Slope	Age : Size/Count	Age : Sum	Age : Sum of Squares	Age : Mean	Age : Median	Age : Mode	Age : Mid Range	
M	ATA	Normal	N	Up	64.0	3028.0	4485.75	47.3125	48.5	52.0, 54.0	45.5	
Age : Standard Deviation	Age : Variance	Age : Range	Age : Mean Absolute Deviation	Age : Minimum	Age : Maximum	Age : Quartiles - Inclusive	Age : Quartiles - Exclusive	Age : Interquartile Range - Inclusive	Age : Interquartile Range - Exclusive			
8.4381503	71.2023810	33.0	7.1894531	29.0	62.0	Q1: 41.000000, Q2: 48.500000, Q3: 54.000000	Q1: 41.000000, Q2: 48.500000, Q3: 54.000000	13.0	13.0			
Age : Coefficient Of Variation	Age : Relative Standard Deviation	Age : Root Mean Square	Age : Standard Error Of Mean	Age : Skewness	Age : Kurtosis	Age : Kurtosis Excess	Age : Shapiro-Wilk: W	Age : Shapiro-Wilk: p-value	Age : Shapiro-Wilk: Prob	Age : Shapiro-Wilk: Decision	Age : Anderson-Darling: A²	
0.1783493	17.8349280	48.047503	1.0547688	-0.3193837	2.2532517	-0.7467483	0.9660686	0.0754195	0.9245805	Fail to Reject Null Hypothesis	0.1169780	
Age : Anderson-Darling: p-value	Age : Anderson-Darling: Prob	Age : Anderson-Darling: Decision	Age : Kolmogorov-Smirnov: D	Age : Kolmogorov-Smirnov: p-value	Age : Kolmogorov-Smirnov: Prob	Age : Kolmogorov-Smirnov: Decision						
0.3199129	0.6800871	Fail to Reject Null Hypothesis	0.7290598	0.0545066	0.9454934	Fail to Reject Null Hypothesis						
RestingBP : Quartiles - Inclusive	RestingBP : Quartiles - Exclusive	RestingBP : Interquartile Range - Inclusive	RestingBP : Interquartile Range - Exclusive	RestingBP : Outliers - Inclusive	RestingBP : Outliers - Exclusive	RestingBP : Coefficient Of Variation	RestingBP : Relative Standard Deviation	RestingBP : Root Mean Square	RestingBP : Standard Error Of Mean	RestingBP : Skewness		
Q1: 120.000000, Q2: 129.000000, Q3: 140.000000	Q1: 120.000000, Q2: 129.000000, Q3: 140.000000	20.0	20.0	190.0	190.0	0.1233311	12.3331060	129.820489	1.9865454	0.8961159		
RestingBP : Kurtosis	RestingBP : Kurtosis Excess	RestingBP : Shapiro-Wilk: W	RestingBP : Shapiro-Wilk: p-value	RestingBP : Shapiro-Wilk: Prob	RestingBP : Shapiro-Wilk: Decision	RestingBP : Anderson-Darling: A²	RestingBP : Anderson-Darling: p-value	RestingBP : Anderson-Darling: Prob	RestingBP : Anderson-Darling: Decision			
5.6138602	2.6138602	0.9245611	0.0007749	0.9992251	Reject Null Hypothesis	0.1636062	0.0577612	0.9422388	Fail to Reject Null Hypothesis			
RestingBP : Size/Count	RestingBP : Sum	RestingBP : Sum of Squares	RestingBP : Mean	RestingBP : Median	RestingBP : Mode	RestingBP : Mid Range	RestingBP : Standard Deviation	RestingBP : Variance	RestingBP : Range	RestingBP : Mean Absolute Deviation	RestingBP : Minimum	RestingBP : Maximum
64.0	8247.0	15911.734375	128.859375	129.0	120.0	144.0	15.8923633	252.5672123	92.0	11.796875	98.0	190.0

To get the plots we must follow the next path:

Data transformation → Data manipulation → Select Column

Then you get the below window:

Select Column(s)

☒ Select Columns Manually

☐ Select Matching Columns from Node

Select Node

☐ Select Columns Manually and from Node

Select Node

Excluded Columns

- Col2 -- Age
- Col3 -- Sex
- Col4 -- ChestPainType
- Col5 -- RestingBP
- Col6 -- Cholesterol
- Col7 -- FastingBS
- Col8 -- RestingECG
- Col9 -- MaxHR
- Col10 -- ExerciseAngina
- Col11 -- Oldpeak
- Col12 -- ST_Slope

Included Columns

Execute Cancel

Select the desired variables from the **Excluded Columns** list and move them to the **Included Columns** list. Once the required columns have been selected, click **Execute**.

Next, create a new tab by clicking the "+" button and give it an appropriate name. Then, right-click on the new tab and select **Import from Spreadsheet**. Click **Execute** again. The following window will appear:

● Action → ● Age dataset source

	Col1	Col2 (I)	Col3	Col4	Col5	Col6
User Header	User Row ID	Age				
1		40				
2		49				
3		37				
4		48				
5		54				
6		39				
7		45				
8		54				
9		37				
10		48				
11		37				
12		58				
13		39				
14		49				
15		42				
16		54				
17		38				

Next, follow the same path as before:

Statistics → Basic Statistics → Descriptive Statistics

Select the type of plot you want to generate and click **Execute**.

The examples below show three commonly used descriptive plots:

- **Q-Q Plot**
- **Histogram**
- **Boxplot** (with **Age** used as the grouping factor)



