



Heart Failure Prediction

The purpose of this research paper is to provide a comprehensive, step-by-step guide for conducting a complete statistical workflow using the Isalos Analytics Platform. Specifically, it demonstrates how to perform a **Multivariate Analysis of Covariance (MANCOVA)** to evaluate the effects of categorical risk factors on multiple correlated cardiovascular outcomes while controlling for the potential confounding effect of age.

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.

MANCOVA extends ANCOVA by simultaneously modelling **multiple continuous dependent variables**, allowing us to assess how categorical factors and a continuous covariate jointly influence a set of interrelated cardiovascular outcomes.

For the purposes of this MANCOVA analysis, variables are classified as follows:

- **Dependent Variables (Responses):**
 - *Resting BP* (Resting Blood Pressure)
 - *Cholesterol* (Serum Cholesterol)

- *Max HR (Maximum Heart Rate)*
(These are continuous outcome measures to be predicted simultaneously).

- **Factors (Categorical Predictors):**

- *Sex (e.g., Male, Female)*
- *Chest Pain Type (e.g., ATA, NAP, ASY, TA)*
(These categorical variables represent demographic/clinical groups whose multivariate effect is of primary interest).

- **Covariate (Continuous Predictor):**

- *Age (in years)*
(This continuous variable is controlled for to remove its confounding influence on the dependent variables).

Step 1: Access the MANCOVA function

Navigate to the MANCOVA analysis module through the main menu by choosing the following path:

Statistics → Analysis of (Co)Variance → MANCOVA

Then, you will see the MANCOVA configuration window:

MANCOVA

?

×

Confidence Level (%)

Double (0,100), Default: -

Sum of Squares for Tests

Adjusted (Type III)

Coding for Factors

(-1, 0, +1)

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

Col13 -- HeartDisease

>

<

>

<

>

<

Dependent Variables

Factors

Specify Reference Levels

Covariates

Formula Type:

☒ Custom
☐ Main Effects
☐ Full Factorial

Specify Custom Model Formula

Formula:

Execute

Cancel

Step 2: Configure MANCOVA parameters

The MANCOVA model can be specified as follows. The model is expressed as:

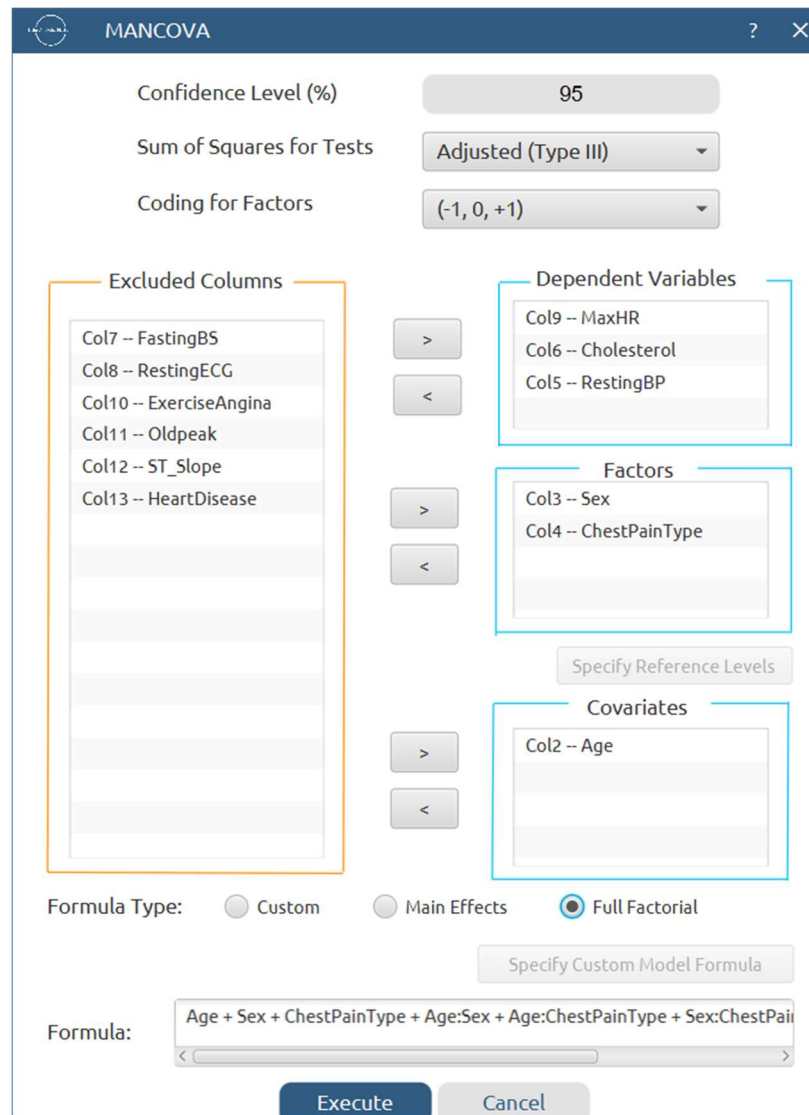
$$[\text{RestingBP}, \text{Cholesterol}, \text{MaxHR}] = \beta_0 + \beta_1(\text{Age}) + \beta_2(\text{Sex}) + \beta_3(\text{ChestPainType}) + \varepsilon$$

This is a multivariate linear model where the three dependent variables are regressed onto the covariate (Age) and the two categorical factors. For this basic demonstration, we focus on the **main-effects model**.

In the MANCOVA configuration window, set the following parameters:

- Confidence Level (%) : 95
- **Dependent Variables:** Move the following to the *Dependent Variables* field:
 - *Resting BP*
 - *Cholesterol*
 - *Max HR*
- **Factors (Categorical Predictors):** Move the following to the *Factors* field:
 - *Sex*
 - *Chest Pain Type*
- **Covariates (Continuous Predictors):** Move the following to the *Covariates* field:
 - *Age*

Then, for **Formula Type**, choose **Full Factorial**. The formula preview will display the multivariate model as shown below:



The image shows a 'MANCOVA' configuration dialog box. At the top, it has a title bar with a question mark and a close button. Below the title bar, there are three settings: 'Confidence Level (%)' set to 95, 'Sum of Squares for Tests' set to 'Adjusted (Type III)', and 'Coding for Factors' set to '(-1, 0, +1)'. The main area is divided into four sections: 'Excluded Columns' (highlighted with an orange box), 'Dependent Variables' (highlighted with a blue box), 'Factors' (highlighted with a blue box), and 'Covariates' (highlighted with a blue box). The 'Excluded Columns' list includes Col7 -- FastingBS, Col8 -- RestingECG, Col10 -- ExerciseAngina, Col11 -- Oldpeak, Col12 -- ST_Slope, and Col13 -- HeartDisease. The 'Dependent Variables' list includes Col9 -- MaxHR, Col6 -- Cholesterol, and Col5 -- RestingBP. The 'Factors' list includes Col3 -- Sex and Col4 -- ChestPainType. The 'Covariates' list includes Col2 -- Age. Below these sections are buttons for '>' and '<' to move items between lists. At the bottom, there are radio buttons for 'Formula Type': 'Custom', 'Main Effects', and 'Full Factorial' (which is selected). Below the radio buttons is a button 'Specify Custom Model Formula'. Below that is a text box for the 'Formula' containing the text 'Age + Sex + ChestPainType + Age:Sex + Age:ChestPainType + Sex:ChestPainType'. At the very bottom are 'Execute' and 'Cancel' buttons.

MANCOVA

Confidence Level (%) 95

Sum of Squares for Tests Adjusted (Type III)

Coding for Factors (-1, 0, +1)

Excluded Columns

- Col7 -- FastingBS
- Col8 -- RestingECG
- Col10 -- ExerciseAngina
- Col11 -- Oldpeak
- Col12 -- ST_Slope
- Col13 -- HeartDisease

Dependent Variables

- Col9 -- MaxHR
- Col6 -- Cholesterol
- Col5 -- RestingBP

Factors

- Col3 -- Sex
- Col4 -- ChestPainType

Specify Reference Levels

Covariates

- Col2 -- Age

Formula Type: ☐ Custom ☐ Main Effects ☒ Full Factorial

Specify Custom Model Formula

Formula: Age + Sex + ChestPainType + Age:Sex + Age:ChestPainType + Sex:ChestPainType

Execute Cancel

After configuring all parameters:

- Review all selections to ensure accuracy.
- Click the **Execute** button to run the MANCOVA analysis.
- Wait for the computation to complete.

Step 3: MANCOVA Results

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

Source	Dependent Variable	DF	Adj SS	Adj MS	F-Value	P-Value	Significant
Age	MaxHR	1	29116.9552506	29116.9552506	60.1152455	0E-7	Yes
	Cholesterol	1	683.2826198	683.2826198	0.0607383	0.8053889	No
	RestingBP	1	10538.9541647	10538.9541647	32.7578599	0E-7	Yes
Sex	MaxHR	1	376.7100253	376.7100253	0.7777604	0.3780619	No
	Cholesterol	1	15146.7460250	15146.7460250	1.3464228	0.2462115	No
	RestingBP	1	3.4136502	3.4136502	0.0106105	0.9179800	No
ChestPainType	MaxHR	3	6373.0590626	2124.3530209	4.3859669	0.0044875	Yes
	Cholesterol	3	1237.1333192	412.3777731	0.0366570	0.9906078	No
	RestingBP	3	161.9233913	53.9744638	0.1677669	0.9181352	No
Age*Sex	MaxHR	1	795.1368804	795.1368804	1.6416500	0.2004285	No
	Cholesterol	1	40042.1216533	40042.1216533	3.5594196	0.0595294	No
	RestingBP	1	22.9428950	22.9428950	0.0713126	0.7894962	No
Age*ChestPainType	MaxHR	3	3381.1519056	1127.0506352	2.3269234	0.0732568	No
	Cholesterol	3	2205.3803347	735.1267782	0.0653468	0.9782120	No
	RestingBP	3	235.9153880	78.6384627	0.2444292	0.8653222	No
Sex*ChestPainType	MaxHR	3	365.4283909	121.8094636	0.2514894	0.8602993	No
	Cholesterol	3	13886.7717012	4628.9239004	0.4114738	0.7447994	No
	RestingBP	3	697.0315305	232.3438435	0.7221862	0.5388434	No

Age*ChestPainType	MaxHR	3	3381.1519056	1127.0506352	2.3269234	0.0732568	No
	Cholesterol	3	2205.3803347	735.1267782	0.0653468	0.9782120	No
	RestingBP	3	235.9153880	78.6384627	0.2444292	0.8653222	No
Sex*ChestPainType	MaxHR	3	365.4283909	121.8094636	0.2514894	0.8602993	No
	Cholesterol	3	13886.7717012	4628.9239004	0.4114738	0.7447994	No
	RestingBP	3	697.0315305	232.3438435	0.7221862	0.5388434	No
Age*Sex*ChestPainType	MaxHR	3	736.6682748	245.5560916	0.5069783	0.6775480	No
	Cholesterol	3	22293.1243371	7431.0414457	0.6605593	0.5764429	No
	RestingBP	3	483.0818047	161.0272682	0.5005154	0.6820098	No
Corrected Model	MaxHR	15	157539.8973976	10502.6598265	21.6839284	0.0	Yes
	Cholesterol	15	824646.8623591	54976.4574906	4.8869608	0E-7	Yes
	RestingBP	15	24129.5997278	1608.6399819	5.0000790	0E-7	Yes
Error	MaxHR	902	436885.7420359	484.3522639			
	Cholesterol	902	10147158.2574666	11249.6211280			
	RestingBP	902	290194.0691175	321.7229148			
Corrected Total	MaxHR	917	594425.6394336				
	Cholesterol	917	10971805.1198257				
	RestingBP	917	314323.6688453				

Step 4 : Residual Analysis & Assumption's checks

Choose

Statistics → Analysis of (Co)Variance → MANCOVA

and choose

Residual Analysis & Assumption's Checks.

You get the following window:


Residual Analysis & Assumption's Checks
?
×

Normality of Residuals

Statistical Tests

☐ Shapiro-Wilk Test

☐ Anderson-Darling Test

☐ Kolmogorov-Smirnov Test

Plots

☐ Normal Q-Q Plots

☐ Histogram of Residuals

☐ Density Plot of Residuals

Homogeneity of Variance (Homoscedasticity)

Statistical Tests

☐ Levene's Test

☐ Brown-Forsythe Test

☐ Bartlett's Test (for normal data)

☐ Box M's Test

Plots

☐ Residuals vs Fitted Values

☐ Residuals vs Dependent Variable

☐ Boxplots by Group

☐ Violin Plot by Group

Residual Diagnostics

Plots

☐ Residuals vs Observation Order

Model Adequacy

Plots

☐ Residuals vs Fitted values

☐ Residuals vs Each Factors

☐ Residuals vs Covariates

Save

Cancel

then choose the tests you want. For example, choose the following:


Residual Analysis & Assumption's Checks
?
×

Normality of Residuals

Statistical Tests

☒ Shapiro-Wilk Test

☒ Anderson-Darling Test

☒ Kolmogorov-Smirnov Test

Plots

☒ Normal Q-Q Plots

☒ Histogram of Residuals

☒ Density Plot of Residuals

Homogeneity of Variance (Homoscedasticity)

Statistical Tests

☒ Levene's Test

☒ Brown-Forsythe Test

☒ Bartlett's Test (for normal data)

☒ Box M's Test

Plots

☒ Residuals vs Fitted Values

☒ Residuals vs Dependent Variable

☒ Boxplots by Group

☒ Violin Plot by Group

Residual Diagnostics

Plots

☒ Residuals vs Observation Order

Model Adequacy

Plots

☒ Residuals vs Fitted values

☒ Residuals vs Each Factors

☒ Residuals vs Covariates

Save

Cancel

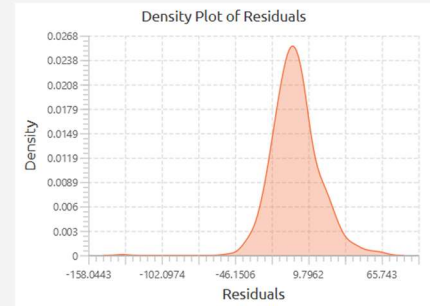
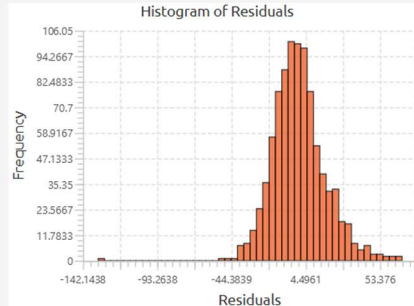
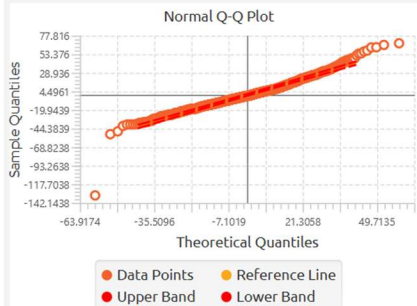
After that, click execute and you get the results (some of them below):

MANCOVA - RestingBP

Normality of Residuals

Statistical Tests

Normality Test	Statistic	Prob	P-value	Decision
Shapiro Wilk	0.965	1.000	0.000	Reject Null Hypothesis
Anderson Darling	4.745	1.000	0.000	Reject Null Hypothesis
Kolmogorov Smirnov	0.063	1.000	0.000	Reject Null Hypothesis



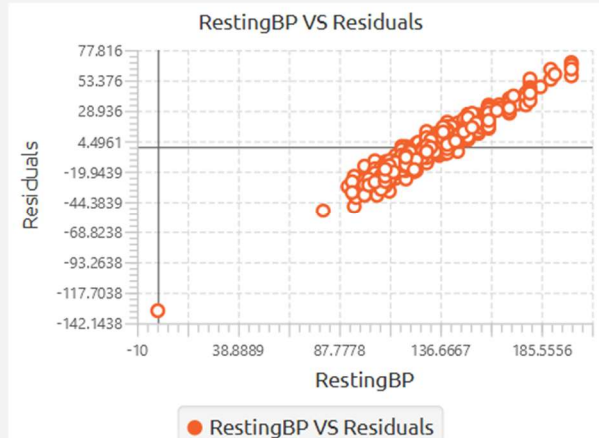
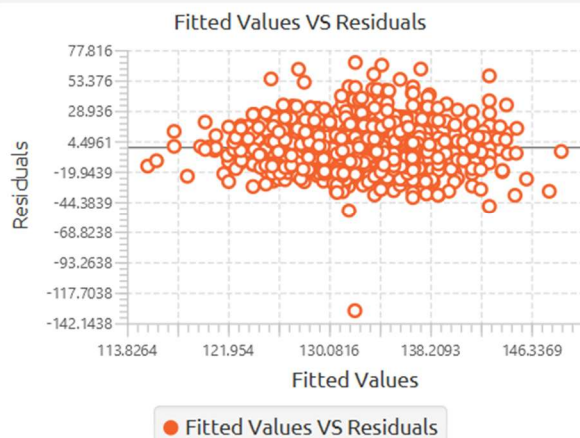
Homogeneity of Variance

Statistical Test (Univariate)

Test	F-statistic	df1	df2	P-value	Decision
Levene	1.295	7	910	0.250	Fail to Reject Null Hypothesis
Brown - Forsythe	1.158	7	910	0.325	Fail to Reject Null Hypothesis
Bartlett	17.414	7		0.015	Reject Null Hypothesis

Statistical Test (Multivariate)

Test	Box's M	F-statistic	df1	df2	P-value	Decision
Box's M	134.574	3.094	42	18785	0.000	Reject Null Hypothesis

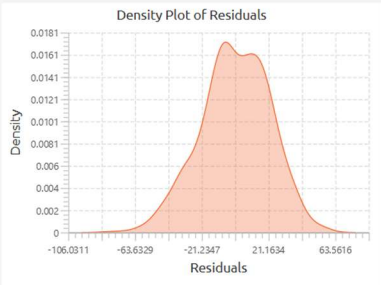
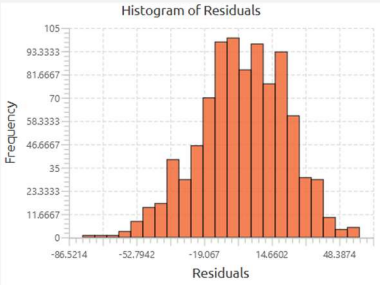
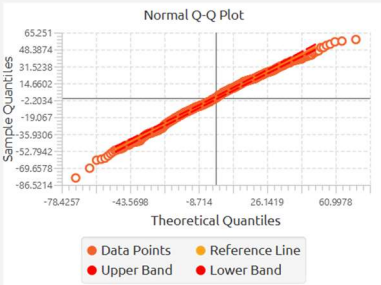




MANCOVA - MaxHR

Normality of Residuals

Statistical Tests				
Normality Test	Statistic	Prob	P-value	Decision
Shapiro Wilk	0.995	0.997	0.003	Reject Null Hypothesis
Anderson Darling	1.304	0.998	0.002	Reject Null Hypothesis
Kolmogorov Smirnov	0.033	0.737	0.263	Fail to Reject Null Hypothesis



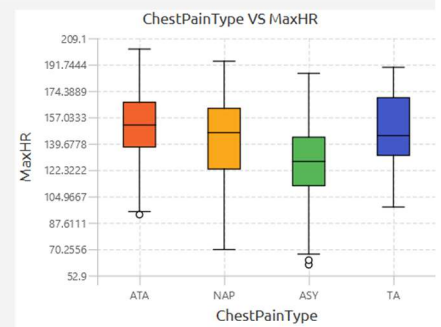
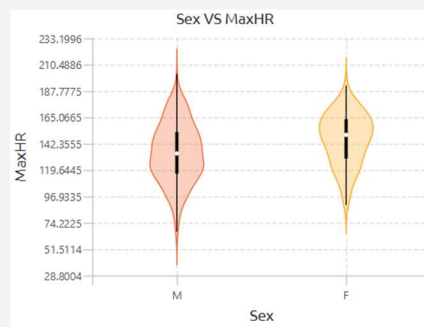
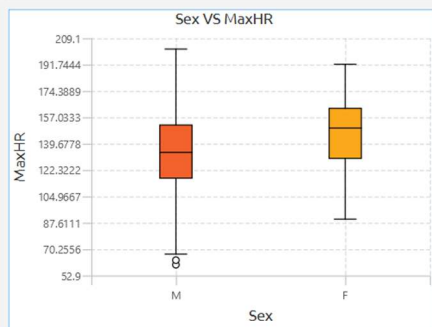
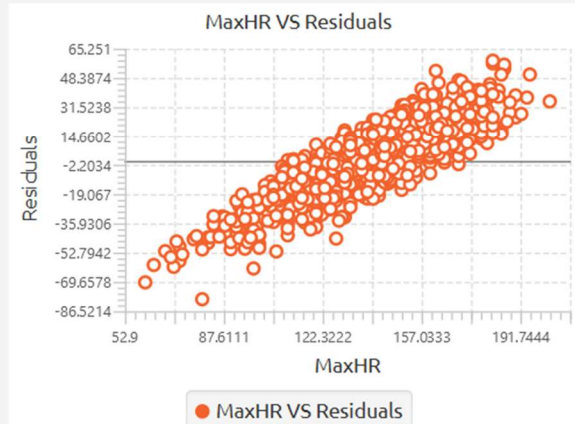
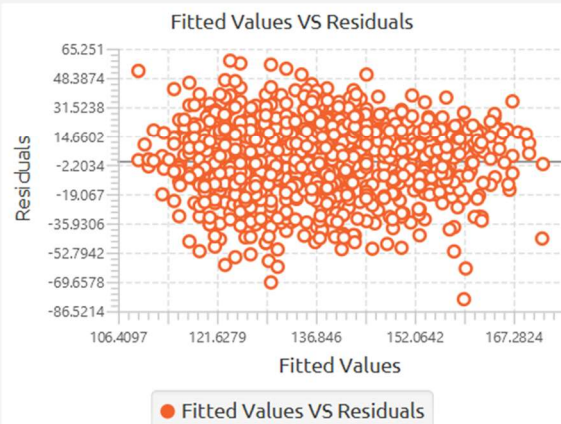
Homogeneity of Variance

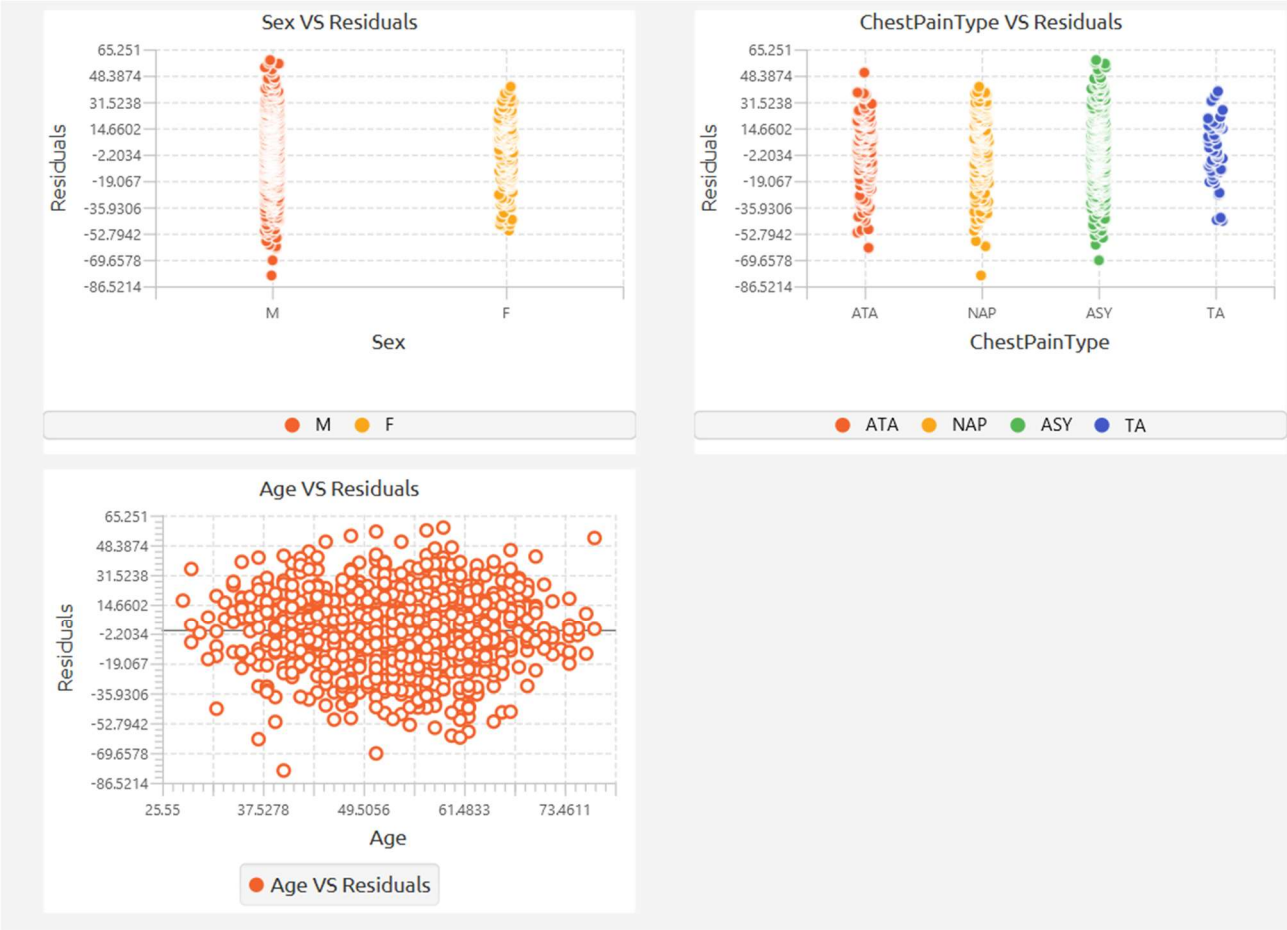
Statistical Test (Univariate)

Test	F-statistic	df1	df2	P-value	Decision
Levene	1.695	7	910	0.107	Fail to Reject Null Hypothesis
Brown - Forsythe	1.661	7	910	0.115	Fail to Reject Null Hypothesis
Bartlett	10.941	7		0.141	Fail to Reject Null Hypothesis

Statistical Test (Multivariate)

Test	Box's M	F-statistic	df1	df2	P-value	Decision
Box's M	134.574	3.094	42	18785	0.000	Reject Null Hypothesis

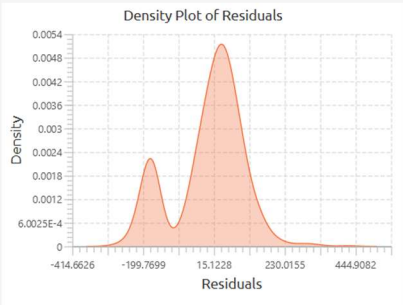
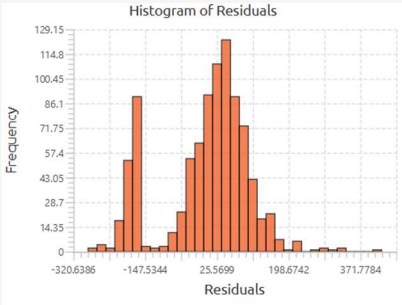
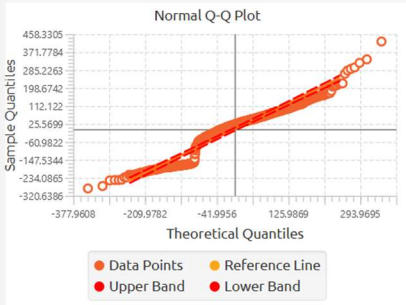




MANCOVA - Cholesterol

Normality of Residuals

Statistical Tests				
Normality Test	Statistic	Prob	P-value	Decision
Shapiro Wilk	0.931	1.000	0.000	Reject Null Hypothesis
Anderson Darling	27.745	1.000	0.000	Reject Null Hypothesis
Kolmogorov Smirnov	0.120	1.000	0.000	Reject Null Hypothesis



Homogeneity of Variance

Statistical Test (Univariate)

Test	F-statistic	df1	df2	P-value	Decision
Levene	16.162	7	910	0.000	Reject Null Hypothesis
Brown - Forsythe	8.829	7	910	0.000	Reject Null Hypothesis
Bartlett	83.593	7		0.000	Reject Null Hypothesis

Statistical Test (Multivariate)

Test	Box's M	F-statistic	df1	df2	P-value	Decision
Box's M	134.574	3.094	42	18785	0.000	Reject Null Hypothesis

