

Lab1 - Data Profiling

claudia.antunes@tecnico.ulisboa.pt [Switch account](#)



Your email will be recorded when you submit this form

* Required

Dataset NYC collisions

Set 1 - Data Dimensionality *

	0 - None	1 - Low	2 - Average	3 - High
Nr Records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nr Variables	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Variables types	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Missing values	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Set 1 - Data granularity *

	0 - None	1 - Low	2 - Average	3 - High	Non Applicable
Numeric variables (dif nr bins)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Symbolic variables (taxonomies)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dates (time hierarchy)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Set 1 - Data Distribution - numeric *

	0 - None	1 - Low	2 - Average	3 - High	Non Applicable
Single boxplot	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Boxplots	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Outliers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Histograms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Probabilities distribution	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Set 1 - Data Distribution - symbolic *

	0 - None	1 - Low	2 - Average	3 - High	Non Applicable
Histograms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Class distribution	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Set 1 - Data Sparsity *

	0 - None	1 - Low	2 - Average	3 - High
Scatter-plots for all pairs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scatter-plots vs class	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Set 1 - Correlation *

	0 - None	1 - Low	2 - Average	3 - High
Correlation between all pairs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correlation with class	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Back

Next

Clear form

This form was created inside of Instituto Superior Técnico. [Report Abuse](#)

Google Forms



Lab1 - Data Profiling

claudia.antunes@tecnico.ulisboa.pt [Switch account](#)



Your email will be recorded when you submit this form

* Required

Dataset AirQuality

Set 2 - Data Dimensionality *

	0 - None	1 - Low	2 - Average	3 - High
Nr Records	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nr Variables	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Variables types	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Missing values	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Set 2 - Data granularity *

	0 - None	1 - Low	2 - Average	3 - High	Non Applicable
Numeric variables (dif nr bins)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Symbolic variables (taxonomies)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dates (time hierarchy)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Set 2 - Data Distribution - numeric *

	0 - None	1 - Low	2 - Average	3 - High	Non Applicable
Single boxplot	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Boxplots	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Outliers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Histograms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Probabilities distribution	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Set 2 - Data Distribution - symbolic *

	0 - None	1 - Low	2 - Average	Non Applicable
Histograms	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Class distribution	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Set 2 - Data Sparsity *

	0 - None	1 - Low	2 - Average	3 - High
Scatter-plots for all pairs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scatter-plots vs class	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Set 2 - Correlation *

	0 - None	1 - Low	2 - Average	3 - High
Correlation between all pairs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correlation with class	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



[Back](#)

Submit

[Clear form](#)

This form was created inside of Instituto Superior Técnico. [Report Abuse](#)

Google Forms

