



DOUGLAS COLLEGE
COMMERCE AND BUSINESS
ADMINISTRATION

Using *german_credit_data.csv* considering *Risk* as the target, implement the following tasks:

- ✓ Find information about “*Banking in Germany*” from *Wikipedia* by *Pageview* package.
- ✓ Get content of the link <https://analyst-2.ai/analysis/kaggle-german-credit-risk-d0f1/97838d73/> and report the type of the page.
- ✓ Convert *Housing* to factor
- ✓ Convert *Checking Account* to factor
- ✓ Create a **parallel box plot** to show the *Credit Amount* against both *Housing* and *Checking Account*
- ✓ Using *sm.density.compare*, show the **density plot** for *Credit Amount* against different levels (values) of *Housing*.
- ✓ Create a **Decision Tree** while calculating the accuracy and confusion matrix.
- ✓ Create a **Random Forest (No. of trees to 50)** while calculating the accuracy and confusion matrix.

The deliverable for each group is a PDF file as the report with members’ name. Your report should include the following tasks, along with screenshots of the codes and results (in the same screenshot) and the explanation for each attempt respectively.

10 points	Excellent	Good	Needs Improvement
Task 1, 2, 3, 4	<ul style="list-style-type: none"> Getting the correct information from the Web Accurate usage of methods Converting two variables to factor 	All requested parts are not presented.	Not enough or readable screenshots and explanations.
(2.5 Points)	(2.5 Point)	(1.25-2.25 Point)	(0-1 Point)
Parallel Box Plot	<ul style="list-style-type: none"> Correct conversion of the variables to factor. Draw the parallel box plot for comparison. 	All requested parts are not presented.	Not enough or readable screenshots and explanations.
(2.5 Points)	(2.5 Point)	(1.25-2.25 Point)	(0-1 Point)
Density Plot	<ul style="list-style-type: none"> Correct conversion of the variables to factor. Draw the density plot for comparison. 	Not enough analyzing and comparisons.	Not enough or readable screenshots and explanations.
(2.5 Points)	(2.5 Point)	(1.25-2.25 Point)	(0-1 Point)
Decision Tee and Random Forest	<ul style="list-style-type: none"> Creating the Random forest Accuracy Confusion matrix. 	All requested parts are not presented.	Not enough or readable screenshots and explanations.
(2.5 Points)	(2.5 Point)	(1.25-2.25 Point)	(0-1 Point)