

Project - 3:

Marked by: Aidan

Name: Ana Cardoso Cachopo

Points: 10/17

Identify: Write a detailed problem statement + Describe Goals & Criteria for Success

2/3 Idea is to give clear metrics by which to judge your model, eg r^2 etc. These should be defined at the start so stakeholders are in agreement.

Acquire: Import with Pandas & create at least one exploratory plot

2/2 Yes good

Parse: State the risks & assumptions of your data

0/2 Didn't see this. That's about being clear with stakeholders issues with the data and underlying assumptions for any model building (eg you are only using a subset of the available data, etc)

Mine: Create necessary columns + Clean/format/process data

2/2 Yes looks good

Refine: Determine outliers, distribution, correlation, statistical analysis

2/2 Good but perhaps give some findings and implications for model building, eg feature selection?

+handle outliers

0/1

Model: Create at least one linear model; calculate probabilities & evaluate model, perform cross-validation, describe bias-variance tradeoff

2/2 Ok good, you can also try cross validation and would be good to have some conclusions and discussion on model variations

+Challenging Bonus: perform Ridge and Lasso regularization and explain what they do

0/1

Present: Create executive summary, present findings, summarize conclusions/recommendations, tailored for your audience (based on your scenario)

0/2

+Bonus: Include additional data & improvement suggestions

0/1

Bonus! Present: Create blog post summary

0/3 I guess you ran out of time on these last parts.