

Python data science

Hamish Gibbs

Data science

What is data science: “Extract meaningful insights from data”.
meaningful is important - use the tools of programming /
statistics to create meaning. You exercise a lot of judgement.

Data science workflow

Not just machine learning - most work is data preparation, data transformation, method selection, communicating results.

Python data science tools

how these line up with the tasks above

R equivalents

pandas (comparison with R), matplotlib (comparison with ggplot) R corollaries” - how these line up with the tools mentioned above