Preparation Material:

Programming

- Kaggle learn: courses 'Intro to programming', 'Python' and 'Pandas'.

- https://automatetheboringstuff.com/ (This is primarily for those who want to learn more about Python)

Probability and Statistics

- Probability distribution functions: https://www.youtube.com/watch?v=YXLVjCKVP7U

- Binomial distribution: https://www.youtube.com/watch?v=e04_wUoscBU

- Normal distribution: https://www.youtube.com/watch?v=RKdB1d5-OE0

- P-values and significance tests: https://www.youtube.com/watch?v=KS6KEWaoOOE

Mathematics

- Matrices: https://www.youtube.com/watch?v=yRwQ7A6jVLk

- Logarithms: https://www.youtube.com/watch?v=kqVpPSzkTYA

- Summation formulas: https://www.youtube.com/watch?v=XJkIaw2e1Pw

- Finding minima and maxima: https://www.youtube.com/watch?v=pvLj1s7SOtk

- Derivatives https://www.youtube.com/watch?v=N2PpRnFqnqY

- Chain rule: https://www.youtube.com/watch?v=0T0QrHO56qg

And from https://www.3blue1brown.com/topics/linear-algebra, at least chapters 1, 2, 3, 4, 6 and 9. This concerns linear algebra and is an important foundation for many DS techniques.

Prerequisites:

Programming skills (in Python)
The student applicant is able to
• set up his/her own development environment to program, test, and run simple Python scripts
• work with integrated development environments such as Python IDE, Jupyter Notebook, PyCharm, VisualStudio Code
• install Python packages
• use Python data types correctly, e.g. numbers, strings, arrays, lists
• apply standard operations on variables
• explain the different types of variables and data structures in Python
• use and write conditional operators e.g. if-else statements, control and loop statements, and own functions
• use modules and packages
• explain the scope of variables
• use basic functionality of NumPy arrays, Pandas data frames, matplotlib plots
• write Python scripts to load and save data, retrieve data from tab-, or comma-delimited files, store data in data objects, visualize data using scatter plots, box plots, histograms.

Statistics
The student applicant
• Is able to use data visualization
• knows mean, median, mode, modal, range, quartile, mean deviation, standard deviation
• knows univariant, bivariant data, scatter plots, outliers, correlation, variance, (co)variance
• knows concepts of probability, relative frequencies, independent and dependent events
• knows concepts of distribution (discrete and continuous)
• knows concepts of sampling, hypothesis testing, statistical tests, and p-value

Mathematics
The student applicant
• knows and correctly applies methods of differential calculus: limits, slope of a function, derivatives, finding minima/maxima, concept of integration
• knows and correctly applies equations and formulas, exponents, logarithm, polynomials, summation and product symbols, function of one or more variables
• knows and correctly applies vector operations and notation
• knows and correctly applies matrix operations and notation