

Table of Contents

Preface	ix
1. Basic Math and Calculus Review	1
Number Theory	2
Order of Operations	3
Variables	5
Functions	6
Summations	11
Exponents	13
Logarithms	16
Euler's Number and Natural Logarithms	18
Euler's Number	18
Natural Logarithms	21
Limits	22
Derivatives	24
Partial Derivatives	28
The Chain Rule	31
Integrals	33
Conclusion	39
Exercises	39
2. Probability	41
Understanding Probability	42
Probability Versus Statistics	43
Probability Math	44
Joint Probabilities	44
Union Probabilities	45
Conditional Probability and Bayes' Theorem	47
Joint and Union Conditional Probabilities	49

Binomial Distribution	51
Beta Distribution	53
Conclusion	60
Exercises	61
3. Descriptive and Inferential Statistics.....	63
What Is Data?	63
Descriptive Versus Inferential Statistics	65
Populations, Samples, and Bias	66
Descriptive Statistics	69
Mean and Weighted Mean	70
Median	71
Mode	73
Variance and Standard Deviation	73
The Normal Distribution	78
The Inverse CDF	85
Z-Scores	87
Inferential Statistics	89
The Central Limit Theorem	89
Confidence Intervals	92
Understanding P-Values	95
Hypothesis Testing	96
The T-Distribution: Dealing with Small Samples	104
Big Data Considerations and the Texas Sharpshooter Fallacy	105
Conclusion	107
Exercises	107
4. Linear Algebra.....	109
What Is a Vector?	110
Adding and Combining Vectors	114
Scaling Vectors	116
Span and Linear Dependence	119
Linear Transformations	121
Basis Vectors	121
Matrix Vector Multiplication	124
Matrix Multiplication	129
Determinants	131
Special Types of Matrices	136
Square Matrix	136
Identity Matrix	136
Inverse Matrix	136
Diagonal Matrix	137
Triangular Matrix	137

Sparse Matrix	138
Systems of Equations and Inverse Matrices	138
Eigenvectors and Eigenvalues	142
Conclusion	145
Exercises	146
5. Linear Regression.....	147
A Basic Linear Regression	149
Residuals and Squared Errors	153
Finding the Best Fit Line	157
Closed Form Equation	157
Inverse Matrix Techniques	158
Gradient Descent	161
Overfitting and Variance	167
Stochastic Gradient Descent	169
The Correlation Coefficient	171
Statistical Significance	174
Coefficient of Determination	179
Standard Error of the Estimate	180
Prediction Intervals	181
Train/Test Splits	185
Multiple Linear Regression	191
Conclusion	191
Exercises	192
6. Logistic Regression and Classification.....	193
Understanding Logistic Regression	193
Performing a Logistic Regression	196
Logistic Function	196
Fitting the Logistic Curve	198
Multivariable Logistic Regression	204
Understanding the Log-Odds	208
R-Squared	211
P-Values	216
Train/Test Splits	218
Confusion Matrices	219
Bayes' Theorem and Classification	222
Receiver Operator Characteristics/Area Under Curve	223
Class Imbalance	225
Conclusion	226
Exercises	226

7. Neural Networks.....	227
When to Use Neural Networks and Deep Learning	228
A Simple Neural Network	229
Activation Functions	231
Forward Propagation	237
Backpropagation	243
Calculating the Weight and Bias Derivatives	243
Stochastic Gradient Descent	248
Using scikit-learn	251
Limitations of Neural Networks and Deep Learning	253
Conclusion	256
Exercise	256
8. Career Advice and the Path Forward.....	257
Redefining Data Science	258
A Brief History of Data Science	260
Finding Your Edge	263
SQL Proficiency	263
Programming Proficiency	266
Data Visualization	269
Knowing Your Industry	270
Productive Learning	272
Practitioner Versus Advisor	272
What to Watch Out For in Data Science Jobs	275
Role Definition	275
Organizational Focus and Buy-In	276
Adequate Resources	278
Reasonable Objectives	279
Competing with Existing Systems	280
A Role Is Not What You Expected	282
Does Your Dream Job Not Exist?	283
Where Do I Go Now?	284
Conclusion	285
A. Supplemental Topics.....	287
B. Exercise Answers.....	309
Index.....	323