

Contents

1	Prelude to The Collaboration	7
1.1	My Approach: The Human Author	7
1.2	Final Thoughts	10
1.3	Summary Poem: The Spark of Collaboration	10
2	The Recipe: The Fundamental Process of Data Scientists	12
2.1	The Spud Gun: From Backyard Curiosity to Competitive Engineering	13
2.2	The Data Science Process: From Questions to Models	14
2.3	Final Thoughts	18
2.4	Summary Poem: The Spud Gun Modeler's Ode	18
3	Aristarchus of Samos: The Redeemed Revolutionary	20
3.1	A Revolution on Hold: The Heliocentric Universe	20
3.2	Aristarchus' Heliocentric Model and His Treatise on Celestial Sizes and Distances	22
3.3	The Measurements	27
3.4	Aristarchus the Data Scientist	33
3.5	Final Thoughts	36
3.6	Summary Poem: Herald of the Sun, Herald of Data Science	36
4	Guo Shoujing and the Shoushi Li—Reconstructing the Solar Calendar	39
4.1	Contrasts: Western and Chinese Approaches to the Heavens	40
4.2	The Observer	42
4.3	From Observations to Forecast	44
4.4	Guo Shoujing, the Data Scientist	52
4.5	Final Thoughts	53
4.6	Summary Poem: Timing the Heavens	53
5	Kepler's Wars: Mars on Earth	56
5.1	Unbelievable	56
5.2	The Battle for Mars	59
5.3	The Struggle	65
5.4	The Legacy	66
5.5	Archimedes' Area, Petiscus' Precision, and Kepler's Calibration	68
5.6	Kepler, the Data Scientist	71
5.7	Final Thoughts	74
5.8	Summary Poem: Kepler's Code	74

6	Flattened: How Gauss and Legendre Conquered the Data	76
6.1	Newton, from his Desk	76
6.2	Lean vs Dumpy	78
6.3	To the Ellipse's Edge	78
6.4	The Meter	80
6.5	Parsing the Data	91
6.6	Gauss, the Data Scientist	94
6.7	Final Thoughts	96
6.8	Summary Poem: The Perfect Fit	96
7	Correlated Cousins: Darwin and Galton	98
7.1	Charles	99
7.2	Francis	103
7.3	Correlation: The Partnership	105
7.4	Galton and Pearson, The Data Scientists	108
7.5	Final Thoughts	110
7.6	Summary Poem: Cousins of Measure	111
8	Upended: Henry Ford and the Industrial Transition	113
8.1	The Purpose	114
8.2	The School of Hard Knocks: Preparation	114
8.3	The Race	116
8.4	Perfecting the Model-T	118
8.5	Mass Production: The Assembly Line	120
8.6	The Cost of Success	123
8.7	The Assembly Line: A Precursor to Modern Data Management and Processing	124
8.8	Henry Ford: The Data Scientist	127
8.9	Final Thoughts	129
8.10	Summary Poem: The Engine and the Algorithm	130
9	AI, Neural Networks and the Connectors	133
9.1	Serendipity, Luigi Galvani (1737–1798)	135
9.2	The Logician, George Boole (1815–1864)	136
9.3	Charles Sherrington (1857-1952): Connecting the Neurons	138
9.4	Claude Shannon (1916–2001) and his Intellectual Empire	138
9.5	Biological Switches: Warren McCulloch and Walter Pitts (1943)	140
9.6	Donald Hebb: But how do we learn?	142
9.7	It's for Real: Hodgkin and Huxley (1949)	143
9.8	Pushing Further Through the Maze: Marvin Minsky (1951)	144
9.9	The Architect: Frank Rosenblatt(1958)	146
9.10	An Essential Midpoint: Widrow and Hoff (1960)	147
9.11	Layering on the Missing Piece: Paul Werblow (1974)	148
9.12	The Other Half	149
9.13	Common Neural Network Architectures	151
9.14	The Penguin Map: RNN's Discovery From Data	159
9.15	The Penguin Map and the the Data Scientist	161
9.16	Final Thoughts	162
9.17	Summary Poem: The Connectors	162

10 The Chat: A Meeting with ChatGPT	165
10.1 Language Processing: From Words to Numbers	166
10.2 From RNNs to Transformers: Maintaining Coherence	167
10.3 The Score	168
10.4 A Superfluous Experiment	171
10.5 The Interview	171
10.6 Summary Poem: The Last Conversation	197
11 Epilogue	199
11.1 Frustration and Relief	199
11.2 No Final Thoughts	203
11.3 Summary Poem: The Mirror and the Maker	203
A Python Code: FFN and RNN for Penguin Map and Bat Map	205
B Further Reading (by chapter)	216