

# Table of Contents

<b>Preface.....</b>	<b>ix</b>
<b>1. Time Series: An Overview and a Quick History.....</b>	<b>1</b>
The History of Time Series in Diverse Applications	2
Medicine as a Time Series Problem	2
Forecasting Weather	6
Forecasting Economic Growth	7
Astronomy	9
Time Series Analysis Takes Off	10
The Origins of Statistical Time Series Analysis	12
The Origins of Machine Learning Time Series Analysis	13
More Resources	13
<b>2. Finding and Wrangling Time Series Data.....</b>	<b>17</b>
Where to Find Time Series Data	18
Prepared Data Sets	18
Found Time Series	25
Retrofitting a Time Series Data Collection from a Collection of Tables	26
A Worked Example: Assembling a Time Series Data Collection	27
Constructing a Found Time Series	33
Timestamping Troubles	35
Whose Timestamp?	35
Guesstimating Timestamps to Make Sense of Data	36
What's a Meaningful Time Scale?	39
Cleaning Your Data	40
Handling Missing Data	40
Upsampling and Downsampling	52
Smoothing Data	55



Seasonal Data	60
Time Zones	63
Preventing Lookahead	67
More Resources	69
<b>3. Exploratory Data Analysis for Time Series.....</b>	<b>73</b>
Familiar Methods	73
Plotting	74
Histograms	77
Scatter Plots	78
Time Series–Specific Exploratory Methods	81
Understanding Stationarity	82
Applying Window Functions	86
Understanding and Identifying Self-Correlation	91
Spurious Correlations	102
Some Useful Visualizations	104
1D Visualizations	104
2D Visualizations	105
3D Visualizations	113
More Resources	117
<b>4. Simulating Time Series Data.....</b>	<b>119</b>
What’s Special About Simulating Time Series?	120
Simulation Versus Forecasting	121
Simulations in Code	121
Doing the Work Yourself	122
Building a Simulation Universe That Runs Itself	128
A Physics Simulation	134
Final Notes on Simulations	140
Statistical Simulations	141
Deep Learning Simulations	141
More Resources	142
<b>5. Storing Temporal Data.....</b>	<b>143</b>
Defining Requirements	145
Live Data Versus Stored Data	146
Database Solutions	148
SQL Versus NoSQL	149
Popular Time Series Database and File Solutions	152
File Solutions	157
NumPy	158
Pandas	158



Standard R Equivalents	158
Xarray	159
More Resources	160
<b>6. Statistical Models for Time Series.....</b>	<b>163</b>
Why Not Use a Linear Regression?	163
Statistical Methods Developed for Time Series	166
Autoregressive Models	166
Moving Average Models	181
Autoregressive Integrated Moving Average Models	186
Vector Autoregression	196
Variations on Statistical Models	201
Advantages and Disadvantages of Statistical Methods for Time Series	203
More Resources	204
<b>7. State Space Models for Time Series.....</b>	<b>207</b>
State Space Models: Pluses and Minuses	209
The Kalman Filter	210
Overview	210
Code for the Kalman Filter	212
Hidden Markov Models	218
How the Model Works	218
How We Fit the Model	220
Fitting an HMM in Code	224
Bayesian Structural Time Series	229
Code for bsts	230
More Resources	235
<b>8. Generating and Selecting Features for a Time Series.....</b>	<b>239</b>
Introductory Example	240
General Considerations When Computing Features	241
The Nature of the Time Series	242
Domain Knowledge	242
External Considerations	243
A Catalog of Places to Find Features for Inspiration	243
Open Source Time Series Feature Generation Libraries	244
Domain-Specific Feature Examples	249
How to Select Features Once You Have Generated Them	252
Concluding Thoughts	255
More Resources	256



<b>9. Machine Learning for Time Series.....</b>	<b>259</b>
Time Series Classification	260
Selecting and Generating Features	260
Decision Tree Methods	264
Clustering	272
Generating Features from the Data	273
Temporally Aware Distance Metrics	280
Clustering Code	285
More Resources	287
<b>10. Deep Learning for Time Series.....</b>	<b>289</b>
Deep Learning Concepts	292
Programming a Neural Network	294
Data, Symbols, Operations, Layers, and Graphs	294
Building a Training Pipeline	298
Inspecting Our Data Set	299
Steps of a Training Pipeline	302
Feed Forward Networks	318
A Simple Example	318
Using an Attention Mechanism to Make Feed Forward	
Networks More Time-Aware	321
CNNs	324
A Simple Convolutional Model	325
Alternative Convolutional Models	327
RNNs	330
Continuing Our Electric Example	332
The Autoencoder Innovation	334
Combination Architectures	335
Summing Up	340
More Resources	341
<b>11. Measuring Error.....</b>	<b>343</b>
The Basics: How to Test Forecasts	344
Model-Specific Considerations for Backtesting	347
When Is Your Forecast Good Enough?	348
Estimating Uncertainty in Your Model with a Simulation	350
Predicting Multiple Steps Ahead	353
Fit Directly to the Horizon of Interest	353
Recursive Approach to Distant Temporal Horizons	354
Multitask Learning Applied to Time Series	354
Model Validation Gotchas	355
More Resources	355



<b>12. Performance Considerations in Fitting and Serving Time Series Models. ....</b>	<b>357</b>
Working with Tools Built for More General Use Cases	358
Models Built for Cross-Sectional Data Don't "Share" Data Across Samples	358
Models That Don't Precompute Create Unnecessary Lag Between	
Measuring Data and Making a Forecast	360
Data Storage Formats: Pluses and Minuses	361
Store Your Data in a Binary Format	361
Preprocess Your Data in a Way That Allows You to "Slide" Over It	362
Modifying Your Analysis to Suit Performance Considerations	362
Using All Your Data Is Not Necessarily Better	363
Complicated Models Don't Always Do Better Enough	363
A Brief Mention of Alternative High-Performance Tools	364
More Resources	365
<b>13. Healthcare Applications. ....</b>	<b>367</b>
Predicting the Flu	367
A Case Study of Flu in One Metropolitan Area	367
What Is State of the Art in Flu Forecasting?	383
Predicting Blood Glucose Levels	384
Data Cleaning and Exploration	385
Generating Features	390
Fitting a Model	396
More Resources	401
<b>14. Financial Applications. ....</b>	<b>403</b>
Obtaining and Exploring Financial Data	404
Preprocessing Financial Data for Deep Learning	410
Adding Quantities of Interest to Our Raw Values	410
Scaling Quantities of Interest Without a Lookahead	411
Formatting Our Data for a Neural Network	413
Building and Training an RNN	416
More Resources	423
<b>15. Time Series for Government. ....</b>	<b>425</b>
Obtaining Governmental Data	426
Exploring Big Time Series Data	428
Upsample and Aggregate the Data as We Iterate Through It	431
Sort the Data	432
Online Statistical Analysis of Time Series Data	436
Remaining Questions	446
Further Improvements	446
More Resources	447



<b>16. Time Series Packages.....</b>	<b>449</b>
Forecasting at Scale	449
Google's Industrial In-house Forecasting	450
Facebook's Open Source Prophet Package	452
Anomaly Detection	457
Twitter's Open Source AnomalyDetection Package	457
Other Time Series Packages	460
More Resources	461
<b>17. Forecasts About Forecasting.....</b>	<b>463</b>
Forecasting as a Service	463
Deep Learning Enhances Probabilistic Possibilities	464
Increasing Importance of Machine Learning Rather Than Statistics	465
Increasing Combination of Statistical and Machine Learning Methodologies	466
More Forecasts for Everyday Life	466
<b>Index.....</b>	<b>469</b>