

# Contents

<b>1</b>	<b>Aggregating Data with Averaging Functions</b>	<b>1</b>
1.1	The Problem in Data: Average Returns	2
1.2	Background Concepts	4
1.2.1	What is a Function?	4
1.2.2	Multivariate Functions and Vectors	7
1.3	The Arithmetic Mean	10
1.3.1	Definition	10
1.3.2	Properties	12
1.4	The Median	14
1.5	The Geometric and Harmonic Means	16
1.5.1	The Geometric Mean	16
1.5.2	More Properties	18
1.5.3	The Harmonic Mean	19
1.6	Averaging Aggregation Functions	21
1.7	Uses of Aggregation Functions	23
1.7.1	Sporting Statistics	24
1.7.2	Simple Forecasting	24
1.7.3	Indices of Diversity and Equity	25
1.8	Summary of Formulas	25
1.9	Practice Questions	26
1.10	R Tutorial	27
1.10.1	Entering Commands in the Console	27
1.10.2	Basic Mathematical Operations	27
1.10.3	Assignment of Variables	28
1.10.4	Assigning a Vector to a Variable	29
1.10.5	Basic Operations on Vectors	29
1.10.6	Existing Functions	30
1.10.7	Creating New Functions	32
1.11	Practice Questions Using R	33
	References	34

<b>2</b>	<b>Transforming Data</b>	37
2.1	The Problem in Data: Multicriteria Evaluation	38
2.1.1	Which is Better: Higher or Lower?	39
2.1.2	Consistent Scales	39
2.1.3	Differences in Distribution	39
2.2	Background Concepts	41
2.2.1	Arrays and Matrices ( $\mathbf{X}$ )	41
2.2.2	Matrix/Array Entries ( $x_{ij}$ )	41
2.2.3	Matrix/Array Rows and Columns ( $\mathbf{x}_i, \mathbf{x}_j$ )	42
2.3	Negations and Utility Transformations	42
2.4	Scaling, Standardization and Normalization	46
2.5	Log and Polynomial Transformations	50
2.6	Piecewise-Linear Transformations	52
2.7	Functions Built from Transformations	56
2.8	Power Means	57
2.9	Quasi-Arithmetic Means	60
2.10	Summary of Formulas	61
2.11	Practice Questions	62
2.12	R Tutorial	63
2.12.1	Replacing Values	63
2.12.2	Arrays and Matrices	64
2.12.3	Reading a Table	65
2.12.4	Transforming Variables	67
2.12.5	Rank-Based Scores	67
2.12.6	Using <code>if ()</code> for Cases	68
2.12.7	Plotting in Two Variables	69
2.12.8	Defining Power Means	71
2.13	Practice Questions Using R	72
	References	72
<b>3</b>	<b>Weighted Averaging</b>	75
3.1	The Problem in Data: Group Decision Making	76
3.2	Background Concepts	78
3.2.1	Regression Parameters	78
3.3	Weighting Vectors	79
3.3.1	Interpreting Weights	81
3.3.2	Example: Welfare Functions	81
3.4	Weighted Power Means	82
3.5	Weighted Medians	84
3.6	Examples of Other Weighting Conventions	85
3.6.1	Simpson's Dominance Index	85
3.6.2	Entropy	86
3.7	The Borda Count	87
3.8	Summary of Formulas	88
3.9	Practice Questions	89

3.10	R Tutorial .....	91
3.10.1	Weighted Arithmetic Means .....	91
3.10.2	Weighted Power Means .....	91
3.10.3	Default Values for Functions .....	92
3.10.4	Weighted Median .....	92
3.10.5	Borda Counts .....	93
3.11	Practice Questions Using R .....	95
	References .....	95
<b>4</b>	<b>Averaging with Interaction .....</b>	<b>97</b>
4.1	The Problem in Data: Supplementary Analytics .....	98
4.2	Background Concepts .....	99
4.2.1	Trimmed and Winsorized Mean .....	99
4.3	Ordered Weighted Averaging .....	101
4.3.1	Definition .....	101
4.3.2	Properties .....	102
4.3.3	Special Cases .....	102
4.3.4	Orness .....	103
4.3.5	Defining Weighting Vectors .....	104
4.4	The Choquet Integral .....	108
4.4.1	Fuzzy Measures .....	108
4.4.2	Definition .....	110
4.4.3	Special Cases .....	111
4.4.4	Calculation .....	111
4.4.5	Examples .....	113
4.4.6	Properties .....	115
4.4.7	The Shapley Value .....	116
4.5	Summary of Formulas .....	117
4.6	Practice Questions .....	118
4.7	R Tutorial .....	120
4.7.1	OWA Operator .....	121
4.7.2	Choquet Integral .....	122
4.7.3	Orness .....	123
4.7.4	Shapley Values .....	125
4.8	Practice Questions Using R .....	125
	References .....	126
<b>5</b>	<b>Fitting Aggregation Functions to Empirical Data .....</b>	<b>129</b>
5.1	The Problem in Data: Recommender Systems .....	130
5.2	Background Concepts .....	134
5.2.1	Optimization and Linear Constraints .....	135
5.2.2	Evaluating and Interpreting a Model's Accuracy .....	139
5.2.3	Flexibility and Overfitting .....	142
5.3	Learning Weights for Aggregation Functions .....	143
5.3.1	Fitting a Weighted Power (or Quasi-Arithmetic) Mean ....	143
5.3.2	Fitting an OWA Function .....	144
5.3.3	Fitting the Choquet Integral .....	145

5.4	Using Aggregation Models for Analysis and Prediction.....	146
5.4.1	Comparing Different Averaging Functions .....	149
5.4.2	Making Inferences About the Importance of Each Variable.....	155
5.4.3	Make Inferences About Tendency Toward Lower or Higher Inputs .....	156
5.4.4	Predicting the Outputs for Unknown/New Data .....	156
5.5	Reliability .....	157
5.5.1	Do Our $y_i$ Values Represent the ‘Ground Truth’? .....	157
5.5.2	Are We Evaluating an Approach or the Model Itself? .....	158
5.6	Conclusions .....	159
5.7	Summary of Formulas .....	160
5.8	Practice Questions Using R .....	161
	References .....	162
<b>6</b>	<b>Solutions .....</b>	<b>163</b>
	Aggregating Data with Averaging Functions: Solutions .....	163
1.9	Practice Questions .....	163
1.10	Practice Questions Using R .....	165
	Transforming Data: Solutions .....	168
2.11	Practice Questions .....	168
2.13	Practice Questions Using R .....	172
	Weighted Aggregation: Solutions .....	176
3.9	Practice Questions .....	176
3.11	Practice Questions Using R .....	180
	Averaging with Interaction: Solutions.....	182
4.6	Practice Questions .....	182
4.7	Practice Questions Using R .....	190
	Fitting Aggregation Functions to Empirical Data: Solutions.....	192
5.6	Practice Questions Using R .....	192
	<b>Index .....</b>	<b>197</b>

An Introduction to Data Analysis using Aggregation  
Functions in R

James, S.

2016, X, 199 p. 29 illus., 20 illus. in color., Hardcover

ISBN: 978-3-319-46761-0