The Science of Sumo: Data, Technique, and Performance in Professional Wrestling

Joseph Babcanec

September 28, 2025

Contents

Preface **xvii**

I Foundations **1**

1 A Short History of Sumo **3**
1.1 Origins and Evolution 3
1.2 The Stable System 3
1.3 Tournament Structure Evolution 4
1.4 Ranking System 4
1.5 Technique Classification 5
1.6 Key Rule Changes Affecting Statistical Analysis 5
1.7 Modern Era Challenges 6
1.8 Data Implications 6

2 Prior Work and Literature Review **9**
2.1 Introduction 9
2.2 Economics and Game Theory 9
2.2.1 The Bubble Effect: Strategic Behavior at 7-7 . . . 9
2.2.2 Tournament Theory and Incentives 9
2.3 Sports Science and Performance 9
2.3.1 Technique Diversity and Success 9
2.3.2 Anthropometric Studies 9
2.4 Medical and Injury Research 9
2.4.1 Injury Epidemiology 9
2.4.2 ACL Injuries in Sumo 9
2.4.3 Metabolic Health 10
2.5 Rating Systems and Prediction 10
2.5.1 Elo-based Approaches 10
2.5.2 Machine Learning Applications 10
2.6 Historical and Cultural Studies 10
2.7 Gaps in the Literature 10

3 Data Infrastructure **11**
3.1 Overview 11
3.2 Primary Data Sources 11
3.2.1 Sumo-API 11
3.2.2 Japan Sumo Association 11
3.2.3 SumoDB 11
3.3 Database Design 11
3.3.1 Schema Architecture 11

3.3.2 Core Tables 11
3.3.3 Derived Tables and Views 11
3.4 Data Collection Pipeline 12
3.4.1 Bulk Import Process 12
3.4.2 Data Validation 12
3.4.3 Update Procedures 12
3.5 Data Quality and Completeness 12
3.5.1 Coverage Analysis 12
3.5.2 Known Limitations 12
3.6 Computational Infrastructure 12
3.7 Reproducibility 12

4 Measurement: Turning Sumo Into Variables 13
4.1 Introduction 13
4.2 Outcome Variables 13
4.2.1 Bout-Level Outcomes 13
4.2.2 Tournament-Level Outcomes 13
4.2.3 Career-Level Outcomes 13
4.3 Wrestler Characteristics 13
4.3.1 Physical Attributes 13
4.3.2 Geographic and Cultural Variables 13
4.3.3 Training Environment 13
4.4 Technique and Style Quantification 14
4.4.1 Kimarite Frequency Vectors 14
4.4.2 Style Classification 14
4.4.3 Technique Diversity Metrics 14
4.4.4 Temporal Evolution 14
4.5 Competition Context 14
4.5.1 Opponent Strength 14
4.5.2 Tournament Position 14
4.5.3 Era and Cohort Effects 14
4.6 Injury and Health Proxies 14
4.6.1 Absence Patterns 14
4.6.2 Performance Degradation 14
4.7 Composite Indices 15
4.8 Measurement Validation 15

5 The Sumo Landscape: Descriptive Analysis 17
5.1 Introduction 17
5.2 Tournament Evolution 17
5.2.1 Structural Changes Over Time 17
5.2.2 Competitive Density 17
5.3 The Kimarite Ecology 17
5.3.1 Technique Frequency Distributions 17
5.3.2 Evolution of Winning Techniques 17

5.3.3 Division-Specific Patterns 17
5.4 Physical Dimensions of Success 18
5.4.1 Anthropometric Distributions 18
5.4.2 The Size-Success Relationship 18
5.4.3 International Comparison 18
5.5 Geographic Patterns 18
5.5.1 Domestic Origins 18
5.5.2 The Foreign Invasion 18
5.5.3 Migration to Tokyo 18
5.6 Stable Dynamics 18
5.6.1 Heya Performance Rankings 18
5.6.2 Training Philosophy Differences 18
5.6.3 The Same-Stable Exclusion Effect 18
5.7 Career Trajectories 19
5.7.1 Typical Career Paths 19
5.7.2 Early Success Indicators 19
5.7.3 Retirement Patterns 19
5.8 Rank Dynamics 19
5.8.1 Promotion and Demotion Flows 19
5.8.2 The Ozeki Bottleneck 19
5.8.3 Yokozuna Dominance 19
5.9 Injury and Absence Patterns 19

II Statistical Analysis 21

6 Statistical Models of Sumo Performance 23
6.1 Introduction 23
6.2 Bout-Level Models 23
6.2.1 Basic Logistic Regression 23
6.2.2 Mixed Effects Models 23
6.2.3 Interaction Effects 23
6.2.4 Model Diagnostics 23
6.3 Tournament-Level Models 23
6.3.1 Kachi-koshi Prediction 23
6.3.2 Win Count Models 23
6.3.3 Rank Change Models 24
6.4 Rating Systems 24
6.4.1 Elo Implementation 24
6.4.2 Glicko and Uncertainty 24
6.4.3 Custom Sumo Ratings 24
6.4.4 Rating Validation 24
6.5 Career Trajectory Models 24
6.5.1 Survival Analysis 24
6.5.2 Multi-State Models 24

6.5.3 Growth Curve Models 24
6.6 Technique Evolution Models 24
6.6.1 Markov Models for Style Changes 24
6.6.2 Diversity and Success 25
6.7 Hierarchical Bayesian Models 25
6.7.1 Wrestler Ability Estimation 25
6.7.2 Stable Effects 25
6.8 Machine Learning Approaches 25
6.8.1 Random Forests 25
6.8.2 Neural Networks 25
6.8.3 Ensemble Methods 25
6.9 Model Comparison 25

7 Causality and Natural Experiments 27
7.1 Introduction 27
7.2 The Fundamental Problem 27
7.2.1 Selection Bias 27
7.2.2 Confounding 27
7.2.3 Reverse Causality 27
7.3 Natural Experiment 1: The 2003 Kosho Abolition 27
7.3.1 Background 27
7.3.2 Identification Strategy 27
7.3.3 Results 27
7.3.4 Robustness Checks 28
7.4 Natural Experiment 2: Foreign Wrestler Quota 28
7.4.1 The 2010 Clarification 28
7.4.2 Regression Discontinuity Design 28
7.4.3 Effects on Competition 28
7.5 The 7-7 Bubble Revisited 28
7.5.1 Replicating Duggan-Levitt 28
7.5.2 Post-Scandal Changes 28
7.5.3 Machine Learning Detection 28
7.6 Instrumental Variables Approaches 28
7.6.1 Birthplace as Instrument 28
7.6.2 Birth Order Effects 28
7.7 Panel Data Methods 29
7.7.1 Fixed Effects Models 29
7.7.2 Dynamic Panel Models 29
7.8 Matching Methods 29
7.8.1 Propensity Score Matching 29
7.8.2 Coarsened Exact Matching 29
7.9 Synthetic Control Methods 29
7.10 Limitations and Caveats 29

III Advanced Topics 31

8 Injuries and Physical Health 33
8.1 Introduction . . . 33
8.2 Injury Epidemiology . . . 33
8.2.1 Prevalence and Incidence . . . 33
8.2.2 Injury Types and Locations . . . 33
8.2.3 Severity Classification . . . 33
8.3 The ACL Crisis . . . 33
8.3.1 Incidence Rates . . . 33
8.3.2 Risk Factors . . . 33
8.3.3 Treatment Decisions . . . 33
8.3.4 Return to Competition . . . 33
8.4 Predictive Models of Injury . . . 34
8.4.1 The Hawkes Process Approach . . . 34
8.4.2 Machine Learning Predictions . . . 34
8.4.3 Individual Risk Profiling . . . 34
8.5 Body Composition and Metabolic Health . . . 34
8.5.1 Extreme Anthropometry . . . 34
8.5.2 Energy Balance . . . 34
8.5.3 Metabolic Syndrome . . . 34
8.5.4 Body Composition Changes . . . 34
8.6 Cardiovascular Health . . . 34
8.6.1 Sleep Apnea . . . 34
8.6.2 Cardiac Events . . . 34
8.6.3 Long-term Cardiovascular Risk . . . 35
8.7 Injury Prevention Strategies . . . 35
8.7.1 Training Modifications . . . 35
8.7.2 Nutritional Interventions . . . 35
8.7.3 Medical Monitoring . . . 35
8.8 The Kosho System and Moral Hazard . . . 35
8.9 Career Impact of Injuries . . . 35
8.9.1 Rank Trajectories . . . 35
8.9.2 Retirement Decisions . . . 35
8.9.3 Financial Implications . . . 35
8.10 Comparative Sports Medicine . . . 35
8.11 Future Directions . . . 36

9 Prediction and Forecasting 37
9.1 Introduction . . . 37
9.2 Short-term Prediction: Next Bout . . . 37
9.2.1 Feature Engineering . . . 37
9.2.2 Model Selection . . . 37
9.2.3 Real-time Prediction . . . 37
9.3 Medium-term: Tournament Outcomes . . . 37

9.3.1 Kachi-koshi Probability 37
9.3.2 Championship Prediction 37
9.3.3 Rank Change Forecasting 37
9.4 Long-term: Career Trajectories 38
9.4.1 Breakthrough Prediction 38
9.4.2 Career Length Forecasting 38
9.4.3 Decline Detection 38
9.5 Rating System Performance 38
9.5.1 Elo Accuracy 38
9.5.2 Glicko Improvements 38
9.5.3 Hybrid Models 38
9.6 Technique-based Prediction 38
9.6.1 Style Matchup Models 38
9.6.2 Adaptation Prediction 38
9.7 Injury Prediction 38
9.7.1 Next Injury Probability 38
9.7.2 Severity Prediction 39
9.8 Ensemble and Meta-learning 39
9.8.1 Model Stacking 39
9.8.2 Online Learning 39
9.9 Prediction Evaluation 39
9.9.1 Scoring Rules 39
9.9.2 Economic Value 39
9.9.3 Comparative Performance 39
9.10 Uncertainty Quantification 39
9.10.1 Prediction Intervals 39
9.10.2 Calibration Analysis 39
9.11 Feature Importance and Interpretability 40
9.11.1 SHAP Values 40
9.11.2 Ablation Studies 40
9.12 Temporal Validation 40
9.13 Case Studies in Prediction 40

10 Case Studies in Sumo Excellence **41**
10.1 Introduction 41
10.2 The Yokozuna Archetypes 41
10.2.1 The Dominant Force: Hakuho 41
10.2.2 The Technician: Kakuryu 41
10.2.3 The Injury Warrior: Terunofuji 41
10.3 Size and Success Stories 41
10.3.1 The Giants: Akebono and Musashimaru 41
10.3.2 The David: Mainoumi 41
10.3.3 The Optimal Build: Asashoryu 42
10.4 International Pioneers 42
10.4.1 The Hawaiian Revolution 42

10.4.2 The Mongolian Dynasty 42
10.4.3 The European Experiment 42
10.5 Technique Specialists 42
10.5.1 The Ultimate Pusher: Akebono 42
10.5.2 The Belt Master: Asashoryu 42
10.5.3 The Innovator: Ura 42
10.6 Career Trajectory Anomalies 42
10.6.1 The Rocket: Onosato 42
10.6.2 The Late Bloomer: Takatoriki 42
10.6.3 The Plateau: Kotooshu 43
10.7 Stable Success Stories 43
10.7.1 The Kokonoe Dynasty 43
10.7.2 The Isegahama Renaissance 43
10.7.3 The Small Stable Miracle 43
10.8 Injury and Redemption 43
10.8.1 The ACL Survivors 43
10.8.2 The Chronic Warriors 43
10.9 The 7-7 Specialists 43
10.9.1 Statistical Anomalies 43
10.9.2 The Bubble Riders 43
10.10 Retirement Decisions 44
10.10.1 The Perfect Exit: Chiyonofuji 44
10.10.2 The Forced Farewell: Asashoryu 44
10.10.3 The Lingerer: Various Examples 44
10.11 Statistical Outliers 44
10.11.1 The Consistency King 44
10.11.2 The Chaos Agent 44
10.11.3 The Specialist 44
10.12 Lessons from the Data 44

11 Where the Data Ends: Future Research 45
11.1 Introduction 45
11.2 Data We Don't Have 45
11.2.1 Training Data 45
11.2.2 Physiological Measurements 45
11.2.3 Psychological Factors 45
11.2.4 Nutritional Details 45
11.2.5 Social Dynamics 45
11.3 Technological Opportunities 45
11.3.1 Computer Vision 45
11.3.2 Wearable Sensors 46
11.3.3 Biometric Monitoring 46
11.3.4 Motion Capture 46
11.4 Advanced Statistical Methods 46
11.4.1 Deep Learning Applications 46

11.4.2 Causal Machine Learning 46
11.4.3 Network Analysis 46
11.4.4 Topological Data Analysis 46
11.5 Unexplored Research Questions 46
11.5.1 Optimal Tournament Design 46
11.5.2 Recruitment Optimization 47
11.5.3 Injury Prevention Programs 47
11.5.4 Post-Career Health 47
11.6 Cross-Sport Comparisons 47
11.6.1 Combat Sports Analysis 47
11.6.2 Size-Dependent Sports 47
11.6.3 Traditional vs. Modern Training 47
11.7 Economic Research 47
11.7.1 Labor Market Analysis 47
11.7.2 Tournament Incentives 47
11.7.3 Stable Economics 47
11.8 Cultural and Social Research 48
11.8.1 Globalization Effects 48
11.8.2 Media and Public Perception 48
11.8.3 Gender and Sumo 48
11.9 Policy Implications 48
11.9.1 Health and Safety Regulations 48
11.9.2 Competition Structure 48
11.9.3 International Development 48
11.10 Methodological Advances Needed 48
11.10.1 Missing Data Methods 48
11.10.2 Multilevel Modeling 48
11.10.3 Computational Challenges 48
11.11 Call for Collaboration 49
11.12 The Next Decade 49

IV Appendices 51

Data Dictionary 53
.1 Database Schema 53
.1.1 wrestlers Table 53
.1.2 tournaments Table 53
.1.3 bouts Table 53
.1.4 banzuke Table 54
.2 Variable Definitions 54
.2.1 Derived Variables 54
.3 Division Codes 55
.4 Missing Data Codes 55
.5 Data Quality Flags 55

Complete Kimarite Catalog **57**
.6 Introduction . . . 57
.7 Basic Techniques (Kihon-waza) . . . 57
.7.1 Push/Thrust Techniques (Oshi-waza) . . . 57
.7.2 Grappling Techniques (Yotsu-waza) . . . 57
.8 Special Techniques (Tokushu-waza) . . . 57
.8.1 Leg Techniques (Ashi-waza) . . . 57
.8.2 Twisting Techniques (Hineri-waza) . . . 58
.8.3 Lifting Techniques (Tsuridashi-waza) . . . 58
.9 Rare Techniques . . . 58
.9.1 Historical Usage . . . 58
.9.2 Modern Rarities . . . 58
.9.3 One-Off Techniques . . . 58
.10 Non-Techniques (Hi-kimarite) . . . 58
.11 Technical Analysis . . . 58
.11.1 Technique Effectiveness . . . 58
.11.2 Evolution Over Time . . . 58
.11.3 Style Correlations . . . 59
.12 Disputed Decisions . . . 59
.13 Training and Teaching . . . 59

Code Repository and Reproducibility **61**
.14 Repository Structure . . . 61
.15 Installation and Setup . . . 61
.15.1 System Requirements . . . 61
.15.2 Database Setup . . . 61
.15.3 Analysis Environment . . . 61
.16 Data Collection Scripts . . . 61
.16.1 bulk_import.py . . . 61
.16.2 API Integration . . . 61
.16.3 Data Validation . . . 62
.17 Analysis Code . . . 62
.17.1 Statistical Models . . . 62
.17.2 Visualization . . . 62
.17.3 Performance Metrics . . . 62
.18 Reproducibility Checklist . . . 62
.19 Data Versioning . . . 62
.20 Testing . . . 62
.20.1 Unit Tests . . . 62
.20.2 Integration Tests . . . 62
.21 Documentation . . . 63
.22 License and Attribution . . . 63

Statistical Methods and Model Details **65**
.23 Model Specifications . . . 65

.23.1 Mixed Effects Logistic Regression 65
.23.2 Survival Models for Career Analysis 65
.23.3 Rating System Mathematics 66
.24 Estimation Methods 66
.24.1 Maximum Likelihood 66
.24.2 Bayesian Estimation 66
.24.3 Cross-Validation Procedures 67
.25 Model Diagnostics 67
.25.1 Residual Analysis 67
.25.2 Influential Observations 67
.25.3 Goodness of Fit 67
.26 Prediction Evaluation Metrics 67
.26.1 Scoring Rules 67
.26.2 Calibration Metrics 68
.27 Computational Details 68
.27.1 Software and Packages 68
.27.2 Convergence Diagnostics 68
.27.3 Computational Complexity 68
.28 Bootstrap and Resampling 68
.28.1 Confidence Intervals 68
.28.2 Bias Correction 68
.29 Multiple Testing Corrections 69
.30 Power Analysis 69
.31 Sensitivity Analysis 69

List of Figures

List of Tables

1.1 Evolution of annual tournament frequency 4

Preface: How to Read This Book

Purpose and Scope

This book represents a comprehensive quantitative analysis of professional sumo wrestling, leveraging modern data science techniques to understand the sport's competitive dynamics, strategic elements, and physical demands. Our approach combines rigorous statistical modeling with domain expertise to illuminate patterns that have shaped sumo from 1958 to the present day.

The scope encompasses:

- Systematic analysis of bout outcomes using mixed-effects models
- Technique ecology and strategic diversity across eras
- Physical anthropometry and its relationship to success
- Injury patterns and career longevity
- Rating systems and predictive modeling
- Natural experiments from policy changes

Data Sources and Reproducibility

All analyses in this book are built on publicly available data sources, primarily:

- **Sumo-API**: Primary structured data source (1958–present) providing rikishi profiles, tournament results, banzuke rankings, and bout outcomes
- **Japan Sumo Association (JSA)**: Official records for verification and supplementary information
- **SumoDB**: Historical cross-validation and extended queries
- **Medical Literature**: PubMed and PMC databases for injury prevalence and health outcomes

Complete code for data acquisition, processing, and analysis is available at our GitHub repository. All random seeds are fixed for reproducibility, and data snapshots are versioned to ensure consistency across analyses.

What "Causality" Does Not Mean Here

This book operates primarily in the realm of observational data analysis. When we discuss relationships between variables—such as the association between technique diversity and win rates—we are describing correlations and predictive patterns, not necessarily causal mechanisms.

True causal inference requires either:

- Randomized controlled experiments (impossible in professional sports)
- Natural experiments with credible identification strategies
- Strong theoretical assumptions that may not hold

We employ quasi-experimental methods where possible, particularly around policy changes like:

- The 2003 abolition of the kōshō (kosho) system (injury leave)
- The 2010 clarification of foreign-born wrestler quotas
- Changes in tournament frequency and structure

However, readers should interpret findings as associations and patterns rather than definitive causal relationships unless explicitly noted otherwise.

Structural Constraints Shaping the Data

Several institutional features of professional sumo create unique data-generating processes:

Foreign-Born Wrestler Cap Each stable (heya (heya)) may have only one foreign-born wrestler at a time. This constraint affects recruitment patterns, career trajectories, and competitive dynamics in ways that must be accounted for in our models.

Same-Stable Non-Competition Rule Wrestlers from the same stable do not face each other in regular tournament bouts (exceptions exist for playoffs). This creates systematic missing data patterns that affect rating calculations and head-to-head analyses.

Promotion/Demotion System The rigid ranking system creates discrete jumps in competition level and incentive structures, particularly around the 7-7 record threshold for maintaining rank.

Tournament Structure Evolution The shift to six tournaments per year in 1958 and other structural changes create era effects that must be controlled for in longitudinal analyses.

How to Use This Book

For the General Reader Focus on Chapters 1–2 for historical context, Chapter 5 for descriptive insights, and Chapter 10 for compelling case studies. Technical sections are clearly marked and can be skipped without losing the narrative thread.

For the Data Scientist Chapters 3–4 detail our data pipeline and feature engineering. Chapters 6–9 contain the core statistical models. Appendix D provides full model specifications and diagnostic tests.

For the Sumo Enthusiast Chapter 5's landscape analysis and Chapter 10's case studies will be of particular interest. Appendix B contains a comprehensive kimarite catalog with historical notes.

For Researchers All data and code are available for replication and extension. We encourage independent verification of our findings and welcome corrections or improvements.

Acknowledgments

This work would not have been possible without the meticulous record-keeping of the Japan Sumo Association and the dedication of the sumo data community, particularly the maintainers of Sumo-API and SumoDB.

Part I

Foundations

Chapter 1

A Short History of Sumo

1.1 Origins and Evolution

Professional sumo wrestling (ōzumō (ozumo)) represents one of the world's oldest continuously practiced sports, with roots extending back over 1,500 years. What began as Shinto harvest rituals evolved through centuries into today's highly structured professional sport, maintaining ceremonial elements while adapting to modern athletic and commercial demands.

The sport's documented history begins in the Nara period (710–794), though archaeological evidence suggests proto-sumo existed centuries earlier. The Nihon Shoki, Japan's second-oldest book of classical history, records a wrestling match in 23 BCE between Nomi no Sukune and Taima no Kuehaya—often cited as sumo's mythological origin.

1.2 The Stable System

The heya (heya) or stable system forms the backbone of professional sumo's organizational structure. These training centers function as:

- Residential facilities where wrestlers live, train, and eat together
- Hierarchical social units with strict seniority systems
- Economic entities that receive JSA subsidies and manage wrestler salaries
- Cultural transmission mechanisms preserving technique and tradition

As of 2024, 44 stables operate under JSA oversight, down from a peak of 54 in the 1990s. Each stable maintains distinct training philosophies, recruitment networks, and technical specializations that measurably affect wrestler development and career outcomes.

Key structural rules governing stables include:

- Maximum of one foreign-born wrestler per stable
- Stable-mates cannot face each other in regular tournament bouts
- Mandatory retirement age of 65 for stablemasters
- Inheritance and branching regulations for stable succession

1.3 Tournament Structure Evolution

The modern six-tournament annual schedule crystallized in 1958, replacing earlier irregular patterns:

Period	Annual Tournaments
Pre-1926	2 (10 days each)
1926–1948	2–4 (variable)
1949–1957	3–5 (11–15 days)
1958–present	6 (15 days each)

Table 1.1: Evolution of annual tournament frequency

Each honbasho (honbasho) (official tournament) follows an identical structure:

- 15 days of competition
- One bout per wrestler per day
- Odd-numbered months (January, March, May, July, September, November)
- Rotating between Tokyo (3), Osaka (1), Nagoya (1), and Fukuoka (1)

This standardization enables consistent statistical analysis across decades while accounting for era-specific effects.

1.4 Ranking System

The banzuke (banzuke) ranking system creates sumo's competitive hierarchy:

Top Division (Makuuchi (makuuchi))

- Yokozuna (grand champion): 0–4 active wrestlers
- Ozeki (champion): 2–5 wrestlers typically
- Sekiwake: Exactly 2 wrestlers minimum
- Komusubi: Exactly 2 wrestlers minimum
- Maegashira: Numbered ranks 1–17 (varies by tournament)

Lower Divisions

- Juryo (salaried): 28 wrestlers
- Makushita: 120 wrestlers
- Sandanme: 180–200 wrestlers

- Jonidan: 200–250 wrestlers
- Jonokuchi: 50–90 wrestlers

Promotion and demotion follow complex algorithms based on:

- Win-loss records (8+ wins typically ensures promotion)
- Current rank
- Available slots in higher divisions
- Special criteria for ozeki and yokozuna

1.5 Technique Classification

The JSA recognizes 82 official winning techniques (kimarite (kimarite)) plus 5 non-techniques (disqualifications, defaults). These broadly categorize into:

Basic Categories

- **Yotsu-zumo** (yotsu-zumo): Belt/grappling techniques (48 techniques)
- **Oshi-zumo** (oshi-zumo): Pushing/thrusting techniques (19 techniques)
- **Nage** (nage): Throwing techniques (included in yotsu)
- **Special techniques**: Leg trips, dodges, rare maneuvers (15 techniques)

Historical kimarite expansion:

- Pre-1960: 48 recognized techniques
- 1960: Expanded to 70 techniques
- 2001: Current 82 techniques established

1.6 Key Rule Changes Affecting Statistical Analysis

Several policy changes create natural experiments for analysis:

1958: Six-Tournament System Standardized annual schedule enabling consistent performance metrics and career progression patterns.

1972: Makushita Tsukedashi Allowed accomplished amateur wrestlers to enter at Makushita 15 or 10, bypassing lower divisions. This created a distinct cohort with different career trajectories.

2003: Abolition of Kōshō System Eliminated injury protection that allowed wrestlers to maintain rank while absent. This fundamentally altered injury incentives and risk-taking behavior:

- Pre-2003: Injured wrestlers could sit out with rank protection
- Post-2003: Any absence results in automatic demotion
- Created measurable changes in injury patterns and career longevity

2010: Foreign Wrestler Definition Clarification Refined rules for naturalized citizens and their stable eligibility, affecting recruitment patterns and competitive balance.

1.7 Modern Era Challenges

Contemporary sumo faces several pressures affecting our data:

Recruitment Crisis New wrestler numbers have declined 40% since 1990, concentrating talent and potentially affecting competitive dynamics.

Internationalization Foreign-born wrestlers now dominate upper ranks despite quotas:

- 67 of 72 yokozuna championships (2004–2024) won by Mongolian wrestlers
- Increasing technical diversity from international styles
- Stable recruitment strategies adapted to one-foreign rule

Match-Fixing Scandals 2011 revelations led to tournament cancellation and 23 forced retirements, creating a structural break in our time series and natural experiment for integrity analysis.

Health and Safety Evolution Increased medical oversight and concussion protocols since 2018 affect injury reporting and potentially bout outcomes.

1.8 Data Implications

These historical and structural elements create several analytical considerations:

1. **Era effects**: Performance metrics must account for rule changes, tournament frequency, and competitive density
2. **Selection bias**: Foreign wrestler quota creates non-random selection into stables

3. **Censoring**: Injury absence patterns changed fundamentally in 2003
4. **Network effects**: Same-stable exclusion creates missing data patterns
5. **Incentive discontinuities**: Promotion thresholds create non-linear effort incentives

Understanding this historical context is essential for interpreting the statistical patterns we uncover in subsequent chapters. The sport's unique blend of ancient tradition and modern athletics creates a rich but complex data environment requiring careful analytical treatment.

Chapter 2

Prior Work and Literature Review

2.1 Introduction

[TO COMPLETE: Overview of existing sumo research across economics, sports science, and medicine]

2.2 Economics and Game Theory

2.2.1 The Bubble Effect: Strategic Behavior at 7-7

[TO COMPLETE: Duggan & Levitt (2002) analysis of match-fixing]

2.2.2 Tournament Theory and Incentives

[TO COMPLETE: Promotion/demotion thresholds and effort provision]

2.3 Sports Science and Performance

2.3.1 Technique Diversity and Success

[TO COMPLETE: University of Hawaii working paper on versatility]

2.3.2 Anthropometric Studies

[TO COMPLETE: Height, weight, BMI relationships with performance]

2.4 Medical and Injury Research

2.4.1 Injury Epidemiology

[TO COMPLETE: Ota et al. (2023) Hawkes process models]

2.4.2 ACL Injuries in Sumo

[TO COMPLETE: 20

2.4.3 Metabolic Health

[TO COMPLETE: Energy expenditure, body composition, metabolic syndrome]

2.5 Rating Systems and Prediction

2.5.1 Elo-based Approaches

[TO COMPLETE: Fan-created rating systems]

2.5.2 Machine Learning Applications

[TO COMPLETE: Neural networks, random forests for outcome prediction]

2.6 Historical and Cultural Studies

[TO COMPLETE: Qualitative research on sumo traditions and evolution]

2.7 Gaps in the Literature

[TO COMPLETE: What this book addresses that prior work has not]

Chapter 3

Data Infrastructure

3.1 Overview

[TO COMPLETE: Introduction to the comprehensive data collection effort]

3.2 Primary Data Sources

3.2.1 Sumo-API

[TO COMPLETE: REST API structure, endpoints, data availability 1958-present]

3.2.2 Japan Sumo Association

[TO COMPLETE: Official records, banzuke publications, wrestler profiles]

3.2.3 SumoDB

[TO COMPLETE: Historical database, query capabilities, cross-validation]

3.3 Database Design

3.3.1 Schema Architecture

[TO COMPLETE: Entity-relationship diagram, normalization decisions]

3.3.2 Core Tables

[TO COMPLETE: wrestlers, tournaments, bouts, banzuke, injuries tables]

3.3.3 Derived Tables and Views

[TO COMPLETE: Aggregations, statistics, performance metrics]

3.4 Data Collection Pipeline

3.4.1 Bulk Import Process

[TO COMPLETE: Python scripts, API rate limiting, error handling]

3.4.2 Data Validation

[TO COMPLETE: Consistency checks, cross-source validation, missing data]

3.4.3 Update Procedures

[TO COMPLETE: Tournament-by-tournament updates, real-time integration]

3.5 Data Quality and Completeness

3.5.1 Coverage Analysis

[TO COMPLETE: Bouts by year, missing tournaments, data gaps]

3.5.2 Known Limitations

[TO COMPLETE: Historical data issues, injury reporting gaps, weight snapshots]

3.6 Computational Infrastructure

[TO COMPLETE: Database performance, indexing strategies, query optimization]

3.7 Reproducibility

[TO COMPLETE: Version control, data snapshots, replication instructions]

Chapter 4

Measurement: Turning Sumo Into Variables

4.1 Introduction

[TO COMPLETE: Philosophy of measurement in sports analytics]

4.2 Outcome Variables

4.2.1 Bout-Level Outcomes

[TO COMPLETE: Win/loss, kimarite, match duration]

4.2.2 Tournament-Level Outcomes

[TO COMPLETE: Kachi-koshi, final record, rank change]

4.2.3 Career-Level Outcomes

[TO COMPLETE: Longevity, highest rank achieved, retirement timing]

4.3 Wrestler Characteristics

4.3.1 Physical Attributes

[TO COMPLETE: Height, weight, BMI, time-varying measurement challenges]

4.3.2 Geographic and Cultural Variables

[TO COMPLETE: Shusshin coding, foreign-born classification, regional effects]

4.3.3 Training Environment

[TO COMPLETE: Heya effects, stablemaster influence, training cohorts]

4.4 Technique and Style Quantification

4.4.1 Kimarite Frequency Vectors

[TO COMPLETE: 82-dimensional technique space, normalization approaches]

4.4.2 Style Classification

[TO COMPLETE: Push vs. grapple spectrum, clustering approaches]

4.4.3 Technique Diversity Metrics

[TO COMPLETE: Shannon entropy, Simpson index, versatility measures]

4.4.4 Temporal Evolution

[TO COMPLETE: Technique changes over career, adaptation patterns]

4.5 Competition Context

4.5.1 Opponent Strength

[TO COMPLETE: Rating differentials, head-to-head records, style matchups]

4.5.2 Tournament Position

[TO COMPLETE: Day effects, pressure situations, clinch scenarios]

4.5.3 Era and Cohort Effects

[TO COMPLETE: Rule changes, competitive density, generational shifts]

4.6 Injury and Health Proxies

4.6.1 Absence Patterns

[TO COMPLETE: Kyujo as injury proxy, severity inference]

4.6.2 Performance Degradation

[TO COMPLETE: Post-injury performance, recovery trajectories]

4.7 Composite Indices

[TO COMPLETE: Overall performance rating, dominance scores, consistency metrics]

4.8 Measurement Validation

[TO COMPLETE: Reliability tests, construct validity, predictive validity]

Chapter 5

The Sumo Landscape: Descriptive Analysis

5.1 Introduction

[TO COMPLETE: Overview of 66 years of professional sumo data]

5.2 Tournament Evolution

5.2.1 Structural Changes Over Time

[TO COMPLETE: Number of wrestlers, tournament format changes, division sizes]

5.2.2 Competitive Density

[TO COMPLETE: Win rate distributions, parity measures, dominance periods]

5.3 The Kimarite Ecology

5.3.1 Technique Frequency Distributions

[TO COMPLETE: Most common techniques by era, rare kimarite occurrences]

5.3.2 Evolution of Winning Techniques

[TO COMPLETE: Time series of technique usage, innovation and obsolescence]

5.3.3 Division-Specific Patterns

[TO COMPLETE: Technique differences across skill levels]

5.4 Physical Dimensions of Success

5.4.1 Anthropometric Distributions

[TO COMPLETE: Height, weight, BMI distributions over time]

5.4.2 The Size-Success Relationship

[TO COMPLETE: Optimal physique analysis, non-linear effects]

5.4.3 International Comparison

[TO COMPLETE: Japanese vs. foreign wrestler characteristics]

5.5 Geographic Patterns

5.5.1 Domestic Origins

[TO COMPLETE: Prefecture-level analysis, regional wrestling styles]

5.5.2 The Foreign Invasion

[TO COMPLETE: Timeline of international participation, country-specific success]

5.5.3 Migration to Tokyo

[TO COMPLETE: Stable recruitment patterns, regional talent pipelines]

5.6 Stable Dynamics

5.6.1 Heya Performance Rankings

[TO COMPLETE: Which stables produce champions, consistency over time]

5.6.2 Training Philosophy Differences

[TO COMPLETE: Technical specialization by stable, size/style preferences]

5.6.3 The Same-Stable Exclusion Effect

[TO COMPLETE: Missing matchups, statistical implications]

5.7 Career Trajectories

5.7.1 Typical Career Paths

[TO COMPLETE: Progression speeds, plateau points, decline patterns]

5.7.2 Early Success Indicators

[TO COMPLETE: Predictors of elite achievement from early career]

5.7.3 Retirement Patterns

[TO COMPLETE: Age, rank, injury influences on retirement]

5.8 Rank Dynamics

5.8.1 Promotion and Demotion Flows

[TO COMPLETE: Transition matrices, expected time at each rank]

5.8.2 The Ozeki Bottleneck

[TO COMPLETE: Difficulty of ozeki promotion, survival analysis]

5.8.3 Yokozuna Dominance

[TO COMPLETE: Championship concentration, reign durations]

5.9 Injury and Absence Patterns

[TO COMPLETE: Kyujo frequency, career impact, recovery success rates]

Part II

Statistical Analysis

Chapter 6

Statistical Models of Sumo Performance

6.1 Introduction

[TO COMPLETE: Modeling philosophy, causal vs. predictive goals]

6.2 Bout-Level Models

6.2.1 Basic Logistic Regression

[TO COMPLETE: P(win) as function of wrestler characteristics]

6.2.2 Mixed Effects Models

[TO COMPLETE: Random effects for wrestlers, tournaments, stables]

$$\text{logit}(P(\text{win}_{ijkt})) = \beta_0 + \beta_1 X_{it} + \beta_2 X_{jt} + \alpha_i + \gamma_j + \delta_k + \epsilon_t \tag{6.1}$$

6.2.3 Interaction Effects

[TO COMPLETE: Style matchups, size interactions, experience differentials]

6.2.4 Model Diagnostics

[TO COMPLETE: Residual analysis, influence diagnostics, cross-validation]

6.3 Tournament-Level Models

6.3.1 Kachi-koshi Prediction

[TO COMPLETE: Binary outcome models, threshold effects at 7-7]

6.3.2 Win Count Models

[TO COMPLETE: Poisson, negative binomial, zero-inflated models]

6.3.3 Rank Change Models

[TO COMPLETE: Ordinal regression for promotion/demotion]

6.4 Rating Systems

6.4.1 Elo Implementation

[TO COMPLETE: K-factor selection, initial ratings, performance]

6.4.2 Glicko and Uncertainty

[TO COMPLETE: Rating deviation, activity bonuses, convergence]

6.4.3 Custom Sumo Ratings

[TO COMPLETE: Incorporating kimarite, injuries, tournament importance]

6.4.4 Rating Validation

[TO COMPLETE: Prediction accuracy, calibration, comparison to banzuke]

6.5 Career Trajectory Models

6.5.1 Survival Analysis

[TO COMPLETE: Time to first makuuchi, career length, retirement hazards]

6.5.2 Multi-State Models

[TO COMPLETE: Transitions between ranks, absorbing states]

6.5.3 Growth Curve Models

[TO COMPLETE: Performance trajectories, peak age, decline rates]

6.6 Technique Evolution Models

6.6.1 Markov Models for Style Changes

[TO COMPLETE: Technique transition matrices, steady states]

6.6.2 Diversity and Success

[TO COMPLETE: Entropy as predictor, optimal versatility]

6.7 Hierarchical Bayesian Models

6.7.1 Wrestler Ability Estimation

[TO COMPLETE: Latent ability parameters, shrinkage, posterior distributions]

6.7.2 Stable Effects

[TO COMPLETE: Hierarchical structure, partial pooling]

6.8 Machine Learning Approaches

6.8.1 Random Forests

[TO COMPLETE: Feature importance, out-of-bag error, interpretability]

6.8.2 Neural Networks

[TO COMPLETE: Architecture, training, performance vs. interpretability]

6.8.3 Ensemble Methods

[TO COMPLETE: Model averaging, stacking, super learners]

6.9 Model Comparison

[TO COMPLETE: AIC/BIC, cross-validation, predictive accuracy metrics]

Chapter 7

Causality and Natural Experiments

7.1 Introduction

[TO COMPLETE: Challenges of causal inference in observational sports data]

7.2 The Fundamental Problem

7.2.1 Selection Bias

[TO COMPLETE: Non-random assignment to stables, weight classes, opponents]

7.2.2 Confounding

[TO COMPLETE: Ability, effort, coaching quality as unobservables]

7.2.3 Reverse Causality

[TO COMPLETE: Does success cause weight gain or vice versa?]

7.3 Natural Experiment 1: The 2003 Kosho Abolition

7.3.1 Background

[TO COMPLETE: Injury protection system, moral hazard concerns]

7.3.2 Identification Strategy

[TO COMPLETE: Difference-in-differences, treated vs. control wrestlers]

7.3.3 Results

[TO COMPLETE: Changes in injury rates, risk-taking, career length]

7.3.4 Robustness Checks

[TO COMPLETE: Parallel trends, placebo tests, sensitivity analysis]

7.4 Natural Experiment 2: Foreign Wrestler Quota

7.4.1 The 2010 Clarification

[TO COMPLETE: Rule change details, affected wrestlers]

7.4.2 Regression Discontinuity Design

[TO COMPLETE: Cutoff exploitation, bandwidth selection, fuzzy RD]

7.4.3 Effects on Competition

[TO COMPLETE: Diversity, success rates, stable composition]

7.5 The 7-7 Bubble Revisited

7.5.1 Replicating Duggan-Levitt

[TO COMPLETE: Updated data, original methodology]

7.5.2 Post-Scandal Changes

[TO COMPLETE: 2011 match-fixing scandal as intervention]

7.5.3 Machine Learning Detection

[TO COMPLETE: Anomaly detection, predictive models of corruption]

7.6 Instrumental Variables Approaches

7.6.1 Birthplace as Instrument

[TO COMPLETE: Distance to Tokyo, local wrestling tradition]

7.6.2 Birth Order Effects

[TO COMPLETE: Family pressure, alternative career options]

7.7 Panel Data Methods

7.7.1 Fixed Effects Models

[TO COMPLETE: Wrestler fixed effects, age-period-cohort models]

7.7.2 Dynamic Panel Models

[TO COMPLETE: Lagged performance, Arellano-Bond estimation]

7.8 Matching Methods

7.8.1 Propensity Score Matching

[TO COMPLETE: Foreign vs. domestic wrestlers, injury effects]

7.8.2 Coarsened Exact Matching

[TO COMPLETE: Stable assignment, physique comparisons]

7.9 Synthetic Control Methods

[TO COMPLETE: Individual wrestler interventions, injury impacts]

7.10 Limitations and Caveats

[TO COMPLETE: External validity, SUTVA violations, interpretation bounds]

Part III

Advanced Topics

Chapter 8

Injuries and Physical Health

8.1 Introduction

[TO COMPLETE: The physical toll of professional sumo]

8.2 Injury Epidemiology

8.2.1 Prevalence and Incidence

[TO COMPLETE: 88% career injury rate, tournament absence patterns]

8.2.2 Injury Types and Locations

[TO COMPLETE: Knee 35-40%, back 25-30%, shoulder 15-20%]

8.2.3 Severity Classification

[TO COMPLETE: Days lost, career-ending vs. minor, chronic vs. acute]

8.3 The ACL Crisis

8.3.1 Incidence Rates

[TO COMPLETE: Primary tears, 20% reinjury rate analysis]

8.3.2 Risk Factors

[TO COMPLETE: Weight, technique, previous injury, age]

8.3.3 Treatment Decisions

[TO COMPLETE: Surgery vs. conservative, rank considerations]

8.3.4 Return to Competition

[TO COMPLETE: Recovery time, performance post-return, re-injury risk]

8.4 Predictive Models of Injury

8.4.1 The Hawkes Process Approach

[TO COMPLETE: Ota et al. (2023) methodology, self-exciting processes]

8.4.2 Machine Learning Predictions

[TO COMPLETE: Feature engineering, model performance, early warning systems]

8.4.3 Individual Risk Profiling

[TO COMPLETE: Personalized injury probability, prevention strategies]

8.5 Body Composition and Metabolic Health

8.5.1 Extreme Anthropometry

[TO COMPLETE: BMI 36.5 average, health implications]

8.5.2 Energy Balance

[TO COMPLETE: 4200-6500 kcal/day expenditure, DLW studies]

8.5.3 Metabolic Syndrome

[TO COMPLETE: Diabetes, hypertension, dyslipidemia prevalence]

8.5.4 Body Composition Changes

[TO COMPLETE: Fat vs. muscle gain, career evolution]

8.6 Cardiovascular Health

8.6.1 Sleep Apnea

[TO COMPLETE: 25-30% prevalence, performance impacts]

8.6.2 Cardiac Events

[TO COMPLETE: OHCA on tournament days, stress factors]

8.6.3 Long-term Cardiovascular Risk

[TO COMPLETE: Post-retirement health, life expectancy]

8.7 Injury Prevention Strategies

8.7.1 Training Modifications

[TO COMPLETE: Load management, technique training, flexibility]

8.7.2 Nutritional Interventions

[TO COMPLETE: Controlled weight gain, supplementation]

8.7.3 Medical Monitoring

[TO COMPLETE: Regular screening, biomarkers, imaging]

8.8 The Kosho System and Moral Hazard

[TO COMPLETE: Pre-2003 injury protection, behavioral changes post-abolition]

8.9 Career Impact of Injuries

8.9.1 Rank Trajectories

[TO COMPLETE: Demotion during recovery, comeback success rates]

8.9.2 Retirement Decisions

[TO COMPLETE: Injury as retirement trigger, survival analysis]

8.9.3 Financial Implications

[TO COMPLETE: Salary loss, career earnings impact]

8.10 Comparative Sports Medicine

[TO COMPLETE: Injury rates vs. other combat sports, unique sumo factors]

8.11 Future Directions

[TO COMPLETE: Wearable technology, biomechanical analysis, prevention programs]

Chapter 9

Prediction and Forecasting

9.1 Introduction

[TO COMPLETE: Predictive modeling in sports, evaluation without gambling]

9.2 Short-term Prediction: Next Bout

9.2.1 Feature Engineering

[TO COMPLETE: Recent form, head-to-head, physical matchup, tournament situation]

9.2.2 Model Selection

[TO COMPLETE: Logistic regression baseline, ensemble methods, neural networks]

9.2.3 Real-time Prediction

[TO COMPLETE: Live tournament predictions, updating with new information]

9.3 Medium-term: Tournament Outcomes

9.3.1 Kachi-koshi Probability

[TO COMPLETE: After day 1, 5, 10 predictions, uncertainty quantification]

9.3.2 Championship Prediction

[TO COMPLETE: Yusho probability, playoff scenarios]

9.3.3 Rank Change Forecasting

[TO COMPLETE: Expected promotion/demotion, banzuke prediction]

9.4 Long-term: Career Trajectories

9.4.1 Breakthrough Prediction

[TO COMPLETE: Time to makuuchi, ozeki probability, yokozuna potential]

9.4.2 Career Length Forecasting

[TO COMPLETE: Expected remaining tournaments, retirement probability]

9.4.3 Decline Detection

[TO COMPLETE: Peak identification, performance degradation patterns]

9.5 Rating System Performance

9.5.1 Elo Accuracy

[TO COMPLETE: Calibration plots, Brier scores, ROC curves]

9.5.2 Glicko Improvements

[TO COMPLETE: Uncertainty benefits, activity adjustments]

9.5.3 Hybrid Models

[TO COMPLETE: Combining ratings with other features]

9.6 Technique-based Prediction

9.6.1 Style Matchup Models

[TO COMPLETE: Push vs. grapple advantages, specific technique success]

9.6.2 Adaptation Prediction

[TO COMPLETE: Technique evolution, counter-strategy development]

9.7 Injury Prediction

9.7.1 Next Injury Probability

[TO COMPLETE: Individual risk scores, temporal dynamics]

9.7.2 Severity Prediction

[TO COMPLETE: Minor vs. career-threatening, recovery time estimation]

9.8 Ensemble and Meta-learning

9.8.1 Model Stacking

[TO COMPLETE: Combining diverse models, weighted averaging]

9.8.2 Online Learning

[TO COMPLETE: Adaptive models, concept drift, tournament-specific tuning]

9.9 Prediction Evaluation

9.9.1 Scoring Rules

[TO COMPLETE: Log loss, Brier score, calibration metrics]

9.9.2 Economic Value

[TO COMPLETE: Decision-theoretic evaluation, without gambling implications]

9.9.3 Comparative Performance

[TO COMPLETE: vs. rankings, vs. public prediction, vs. random]

9.10 Uncertainty Quantification

9.10.1 Prediction Intervals

[TO COMPLETE: Bootstrapping, Bayesian credible intervals]

9.10.2 Calibration Analysis

[TO COMPLETE: Reliability diagrams, isotonic regression]

9.11 Feature Importance and Interpretability

9.11.1 SHAP Values

[TO COMPLETE: Individual prediction explanation, global importance]

9.11.2 Ablation Studies

[TO COMPLETE: Removing features, marginal contribution]

9.12 Temporal Validation

[TO COMPLETE: Rolling window, expanding window, tournament-by-tournament]

9.13 Case Studies in Prediction

[TO COMPLETE: Specific tournament predictions, upset analysis, model failures]

Chapter 10

Case Studies in Sumo Excellence

10.1 Introduction

[TO COMPLETE: Individual stories illuminating statistical patterns]

10.2 The Yokozuna Archetypes

10.2.1 The Dominant Force: Hakuho

[TO COMPLETE: 45 championships, statistical dominance, longevity factors]

10.2.2 The Technician: Kakuryu

[TO COMPLETE: Versatility metrics, adaptation over career]

10.2.3 The Injury Warrior: Terunofuji

[TO COMPLETE: Comeback from Jonidan, injury management, second peak]

10.3 Size and Success Stories

10.3.1 The Giants: Akebono and Musashimaru

[TO COMPLETE: Maximum size advantage, technique limitations, health costs]

10.3.2 The David: Mainoumi

[TO COMPLETE: Success despite size disadvantage, technique diversity necessity]

10.3.3 The Optimal Build: Asashoryu

[TO COMPLETE: Power-to-weight ratio, speed advantage, complete package]

10.4 International Pioneers

10.4.1 The Hawaiian Revolution

[TO COMPLETE: Takamiyama to Akebono, cultural adaptation, training differences]

10.4.2 The Mongolian Dynasty

[TO COMPLETE: Technical innovation, hunger for success, domination statistics]

10.4.3 The European Experiment

[TO COMPLETE: Baruto, Tochinoshin, different body types and styles]

10.5 Technique Specialists

10.5.1 The Ultimate Pusher: Akebono

[TO COMPLETE: Oshi-zumo perfection, kimarite concentration]

10.5.2 The Belt Master: Asashoryu

[TO COMPLETE: Yotsu-zumo excellence, grip strength advantage]

10.5.3 The Innovator: Ura

[TO COMPLETE: Rare technique usage, unpredictability as strategy]

10.6 Career Trajectory Anomalies

10.6.1 The Rocket: Onosato

[TO COMPLETE: Fastest rise to yokozuna, early indicators, sustainability]

10.6.2 The Late Bloomer: Takatoriki

[TO COMPLETE: Delayed success, persistence, peak timing]

10.6.3 The Plateau: Kotooshu

[TO COMPLETE: Ozeki struggles, inability to break through, consistency vs. excellence]

10.7 Stable Success Stories

10.7.1 The Kokonoe Dynasty

[TO COMPLETE: Multiple yokozuna, training methods, recruitment]

10.7.2 The Isegahama Renaissance

[TO COMPLETE: Terunofuji and Harumafuji, Mongolian connection]

10.7.3 The Small Stable Miracle

[TO COMPLETE: Success despite limited resources, individual attention benefits]

10.8 Injury and Redemption

10.8.1 The ACL Survivors

[TO COMPLETE: Successful returns, technique adaptations, career modifications]

10.8.2 The Chronic Warriors

[TO COMPLETE: Managing persistent injuries, strategic absences]

10.9 The 7-7 Specialists

10.9.1 Statistical Anomalies

[TO COMPLETE: Wrestlers with unusual 8-7 frequencies, clutch performance]

10.9.2 The Bubble Riders

[TO COMPLETE: Career patterns around rank thresholds]

10.10 Retirement Decisions

10.10.1 The Perfect Exit: Chiyonofuji

[TO COMPLETE: Timing, legacy preservation, post-career success]

10.10.2 The Forced Farewell: Asashoryu

[TO COMPLETE: Scandal, premature end, what-if scenarios]

10.10.3 The Lingerer: Various Examples

[TO COMPLETE: Staying too long, rank preservation, financial motivations]

10.11 Statistical Outliers

10.11.1 The Consistency King

[TO COMPLETE: Lowest variance in performance, reliability metrics]

10.11.2 The Chaos Agent

[TO COMPLETE: Highest upset rate, giant-killer profile]

10.11.3 The Specialist

[TO COMPLETE: Dominance against specific opponents/styles]

10.12 Lessons from the Data

[TO COMPLETE: Common patterns, success factors, predictive indicators]

Chapter 11

Where the Data Ends: Future Research

11.1 Introduction

[TO COMPLETE: Limitations of current data, unexplored territories]

11.2 Data We Don't Have

11.2.1 Training Data

[TO COMPLETE: Keiko intensity, practice bouts, morning training routines]

11.2.2 Physiological Measurements

[TO COMPLETE: VO2 max, muscle composition, hormone levels, sleep quality]

11.2.3 Psychological Factors

[TO COMPLETE: Mental health, motivation, pressure handling, personality]

11.2.4 Nutritional Details

[TO COMPLETE: Actual caloric intake, meal timing, supplementation]

11.2.5 Social Dynamics

[TO COMPLETE: Stable relationships, mentorship quality, hazing effects]

11.3 Technological Opportunities

11.3.1 Computer Vision

[TO COMPLETE: Automated kimarite classification, stance analysis, movement patterns]

11.3.2 Wearable Sensors

[TO COMPLETE: Accelerometers, force platforms, real-time biomechanics]

11.3.3 Biometric Monitoring

[TO COMPLETE: Heart rate variability, recovery metrics, stress indicators]

11.3.4 Motion Capture

[TO COMPLETE: 3D movement analysis, technique optimization, injury prevention]

11.4 Advanced Statistical Methods

11.4.1 Deep Learning Applications

[TO COMPLETE: Video analysis, pattern recognition, complex interactions]

11.4.2 Causal Machine Learning

[TO COMPLETE: Heterogeneous treatment effects, optimal policy learning]

11.4.3 Network Analysis

[TO COMPLETE: Wrestler relationships, influence propagation, strategic networks]

11.4.4 Topological Data Analysis

[TO COMPLETE: Shape of technique space, performance manifolds]

11.5 Unexplored Research Questions

11.5.1 Optimal Tournament Design

[TO COMPLETE: Schedule effects, fatigue management, fairness optimization]

11.5.2 Recruitment Optimization

[TO COMPLETE: Talent identification, development pathways, resource allocation]

11.5.3 Injury Prevention Programs

[TO COMPLETE: Personalized training, load management, early intervention]

11.5.4 Post-Career Health

[TO COMPLETE: Long-term outcomes, life expectancy, quality of life]

11.6 Cross-Sport Comparisons

11.6.1 Combat Sports Analysis

[TO COMPLETE: MMA, judo, wrestling parallels, technique transfer]

11.6.2 Size-Dependent Sports

[TO COMPLETE: NFL linemen, rugby forwards, weightlifting]

11.6.3 Traditional vs. Modern Training

[TO COMPLETE: Eastern vs. Western approaches, evolution of methods]

11.7 Economic Research

11.7.1 Labor Market Analysis

[TO COMPLETE: Career earnings, outside options, retirement timing]

11.7.2 Tournament Incentives

[TO COMPLETE: Prize structure optimization, effort provision]

11.7.3 Stable Economics

[TO COMPLETE: Resource allocation, recruitment ROI, training investments]

11.8 Cultural and Social Research

11.8.1 Globalization Effects

[TO COMPLETE: Cultural exchange, technique diffusion, fan demographics]

11.8.2 Media and Public Perception

[TO COMPLETE: Narrative effects, star creation, scandal impacts]

11.8.3 Gender and Sumo

[TO COMPLETE: Women's amateur sumo, cultural barriers, future possibilities]

11.9 Policy Implications

11.9.1 Health and Safety Regulations

[TO COMPLETE: Weight limits, concussion protocols, mandatory rest]

11.9.2 Competition Structure

[TO COMPLETE: Division size, tournament frequency, ranking system]

11.9.3 International Development

[TO COMPLETE: Olympic inclusion, global expansion, rule standardization]

11.10 Methodological Advances Needed

11.10.1 Missing Data Methods

[TO COMPLETE: Injury under-reporting, selection bias, censoring]

11.10.2 Multilevel Modeling

[TO COMPLETE: Bout-day-tournament-career hierarchies]

11.10.3 Computational Challenges

[TO COMPLETE: Big data infrastructure, real-time processing, distributed computing]

11.11 Call for Collaboration

[TO COMPLETE: Data sharing protocols, research partnerships, open science]

11.12 The Next Decade

[TO COMPLETE: Predictions for sumo evolution, research priorities, potential breakthroughs]

Part IV

Appendices

Data Dictionary

.1 Database Schema

.1.1 wrestlers Table

Field	Type	Description
id	INTEGER	Unique wrestler identifier
shikona	TEXT	Ring name (may change over career)
real_name	TEXT	Birth name
birth_date	DATE	Date of birth
debut_date	DATE	Professional debut date
retirement_date	DATE	Retirement date (NULL if active)
height_cm	INTEGER	Height in centimeters
weight_kg	INTEGER	Weight in kilograms
shusshin	TEXT	Birthplace/origin
heya	TEXT	Stable affiliation
foreign_born	BOOLEAN	True if born outside Japan

.1.2 tournaments Table

Field	Type	Description
id	TEXT	Tournament ID (YYYYMM format)
year	INTEGER	Tournament year
month	INTEGER	Tournament month (1,3,5,7,9,11)
location	TEXT	City (Tokyo, Osaka, Nagoya, Fukuoka)
start_date	DATE	First day of tournament
end_date	DATE	Final day of tournament

.1.3 bouts Table

Field	Type	Description
id	INTEGER	Unique bout identifier
tournament_id	TEXT	Foreign key to tournaments
day	INTEGER	Day of tournament (1-15)
division	TEXT	Competition division
east_wrestler_id	INTEGER	Eastern wrestler ID
west_wrestler_id	INTEGER	Western wrestler ID

Field	Type	Description
winner_id	INTEGER	Winner's ID
kimarite	TEXT	Winning technique
match_time_seconds	INTEGER	Duration in seconds

.1.4 banzuke Table

Field	Type	Description
tournament_id	TEXT	Foreign key to tournaments
wrestler_id	INTEGER	Foreign key to wrestlers
division	TEXT	Division name
rank	TEXT	Rank designation (Y, O, S, K, M1-17, J1-14)
rank_number	INTEGER	Numeric rank for ordering
east_west	TEXT	East or West designation

.2 Variable Definitions

.2.1 Derived Variables

win_rate: Total wins divided by total bouts (wins + losses)

kachi_koshi: Binary indicator for achieving winning record (8+ wins)

make_koshi: Binary indicator for losing record (8+ losses)

career_length: Years from debut_date to retirement_date or current date

bmi: weight_kg / (height_cm/100)^2

push_ratio: Proportion of wins by pushing techniques

technique_diversity: Shannon entropy of kimarite distribution

kyujo_rate: Proportion of tournaments with absences

.3 Division Codes

Code	Division Name
Y	Yokozuna
O	Ozeki
S	Sekiwake
K	Komusubi
M	Maegashira
J	Juryo
Ms	Makushita
Sd	Sandanme
Jd	Jonidan
Jk	Jonokuchi

.4 Missing Data Codes

Code	Meaning
NULL	Value unknown or not applicable
0	Confirmed zero/absent
-1	Not yet occurred (future event)
-999	Error or invalid data

.5 Data Quality Flags

[TO COMPLETE: Reliability indicators, source verification, imputation flags]

Complete Kimarite Catalog

.6 Introduction

[TO COMPLETE: All 82 official winning techniques plus 5 non-techniques]

.7 Basic Techniques (Kihon-waza)

.7.1 Push/Thrust Techniques (Oshi-waza)

Technique	Japanese	Description	Frequency
oshidashi	oshidashi	Push out	15.2%
oshitaoshi	oshitaoshi	Push down	8.1%
tsukidashi	tsukidashi	Thrust out	3.4%
tsukitaoshi	tsukitaoshi	Thrust down	1.8%
hatakikomi	hatakikomi	Slap down	6.7%

[TO COMPLETE: All push techniques with descriptions and historical frequencies]

.7.2 Grappling Techniques (Yotsu-waza)

Technique	Japanese	Description	Frequency
yorikiri	yorikiri	Force out	32.1%
uwatenage	uwatenage	Overarm throw	4.2%
shitatenage	shitatenage	Underarm throw	2.8%
kotenage	kotenage	Arm throw	1.6%
sukuinage	sukuinage	Scooping throw	1.3%

[TO COMPLETE: All grappling techniques]

.8 Special Techniques (Tokushu-waza)

.8.1 Leg Techniques (Ashi-waza)

[TO COMPLETE: All leg-based techniques]

.8.2 Twisting Techniques (Hineri-waza)

[TO COMPLETE: All twisting/turning techniques]

.8.3 Lifting Techniques (Tsuridashi-waza)

[TO COMPLETE: All lifting techniques]

.9 Rare Techniques

.9.1 Historical Usage

[TO COMPLETE: Techniques that were more common in earlier eras]

.9.2 Modern Rarities

[TO COMPLETE: Techniques rarely seen in contemporary sumo]

.9.3 One-Off Techniques

[TO COMPLETE: Techniques used only a few times in recorded history]

.10 Non-Techniques (Hi-kimarite)

Non-technique	Japanese	Description
hansoku	hansoku	Foul (illegal move)
koshikudake	koshikudake	Collapse without contact
fumidashi	fumidashi	Step out
isamiashi	isamiashi	Overeager step out
tottari	tottari	Arm bar throw

.11 Technical Analysis

.11.1 Technique Effectiveness

[TO COMPLETE: Win rates by technique, situational effectiveness]

.11.2 Evolution Over Time

[TO COMPLETE: How technique frequencies have changed 1958-2024]

.11.3 Style Correlations

[TO COMPLETE: Which techniques correlate with body type, origin, stable]

.12 Disputed Decisions

[TO COMPLETE: Controversial kimarite calls, video review impacts]

.13 Training and Teaching

[TO COMPLETE: Which techniques are taught first, difficulty progression]

Code Repository and Reproducibility

.14 Repository Structure

[TO COMPLETE: GitHub repository layout]

.15 Installation and Setup

.15.1 System Requirements

- Python 3.8+ with packages: pandas, numpy, sqlite3, requests
- SQLite 3.35+
- LaTeX distribution (TeX Live recommended)
- R 4.0+ (optional, for some analyses)

.15.2 Database Setup

```
cd database/
sqlite3 sumo.db < setup.sql
python bulk_import.py
```

.15.3 Analysis Environment

[TO COMPLETE: Jupyter notebook setup, package versions, environment files]

.16 Data Collection Scripts

.16.1 bulk_import.py

[TO COMPLETE: Main data collection script documentation]

.16.2 API Integration

[TO COMPLETE: Sumo-API wrapper functions, rate limiting, error handling]

.16.3 Data Validation

[TO COMPLETE: Consistency checks, cross-source validation]

.17 Analysis Code

.17.1 Statistical Models

[TO COMPLETE: Model fitting scripts, validation procedures]

.17.2 Visualization

[TO COMPLETE: Figure generation code, plot styling]

.17.3 Performance Metrics

[TO COMPLETE: Rating calculation, prediction evaluation]

.18 Reproducibility Checklist

1. Clone repository: `git clone [URL]`
2. Install dependencies: `pip install -r requirements.txt`
3. Run data collection: `python database/bulk_import.py`
4. Execute analyses: `python analysis/run_all.py`
5. Generate figures: `python figures/generate_all.py`
6. Compile book: `pdflatex main.tex`

.19 Data Versioning

[TO COMPLETE: Data snapshot management, version control for large files]

.20 Testing

.20.1 Unit Tests

[TO COMPLETE: Model validation, data integrity tests]

.20.2 Integration Tests

[TO COMPLETE: End-to-end pipeline validation]

.21 Documentation

[TO COMPLETE: Function documentation, API reference, tutorial notebooks]

.22 License and Attribution

[TO COMPLETE: Open source license, data source attribution, citation requirements]

Statistical Methods and Model Details

.23 Model Specifications

.23.1 Mixed Effects Logistic Regression

The primary bout outcome model is specified as:

$$\text{logit}(P(Y_{ijkt} = 1)) = \beta_0 + \boldsymbol{\beta}_1^T \mathbf{X}_{it} + \boldsymbol{\beta}_2^T \mathbf{X}_{jt} \tag{1}$$
$$+ \boldsymbol{\beta}_3^T \mathbf{Z}_{kt} + \alpha_i + \gamma_j + \delta_k + \epsilon_t \tag{2}$$

where:

- $Y_{ijkt} = 1$ if wrestler i beats wrestler j in tournament k on day t
- $\mathbf{X}_{it}$ = wrestler i's characteristics at time t
- $\mathbf{Z}_{kt}$ = tournament/context variables
- $\alpha_i \sim N(0, \sigma_\alpha^2)$ = wrestler random effects
- $\gamma_j \sim N(0, \sigma_\gamma^2)$ = opponent random effects
- $\delta_k \sim N(0, \sigma_\delta^2)$ = tournament random effects
- $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ = day effects

.23.2 Survival Models for Career Analysis

For career longevity analysis, we employ Cox proportional hazards models:

$$h(t|\mathbf{X}) = h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{X}) \tag{3}$$

With competing risks framework for different retirement types:

- Voluntary retirement
- Forced retirement (injury)
- Demotion-induced retirement

.23.3 Rating System Mathematics

Elo System

Updates follow:

$$R_{i,new} = R_{i,old} + K \cdot (S_i - E_i) \tag{4}$$

$$E_i = \frac{1}{1 + 10^{(R_j - R_i)/400}} \tag{5}$$

where K varies by division and experience:

$$K = \begin{cases} 40 & \text{if new wrestler (first 10 tournaments)} \\ 32 & \text{if makuuchi/juryo} \\ 24 & \text{if lower divisions} \end{cases} \tag{6}$$

Glicko System

Rating updates incorporate uncertainty (rating deviation):

$$\phi^2_{new} = \frac{1}{\frac{1}{\phi^2_{old} + \sigma^2} + \frac{1}{v}} \tag{7}$$

$$\mu_{new} = \mu_{old} + \phi^2_{new} \sum_j \frac{g(\phi_j)(s_j - E(\mu, \mu_j, \phi_j))}{\sigma^2} \tag{8}$$

.24 Estimation Methods

.24.1 Maximum Likelihood

For logistic regression models, we maximize:

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[y_i \log p_i + (1 - y_i) \log(1 - p_i) \right] \tag{9}$$

.24.2 Bayesian Estimation

Hierarchical models estimated using MCMC with priors:

$$\boldsymbol{\beta} \sim N(\mathbf{0}, \tau^2 \mathbf{I}) \tag{10}$$

$$\sigma^2_\alpha \sim \text{InvGamma}(a, b) \tag{11}$$

$$\tau^2 \sim \text{InvGamma}(c, d) \tag{12}$$

.24.3 Cross-Validation Procedures

Time series cross-validation using expanding windows:

- Training: Tournaments 1 to t
- Testing: Tournament $t + 1$
- Advance t and repeat

.25 Model Diagnostics

.25.1 Residual Analysis

Pearson residuals for logistic models:

$$r_i = \frac{y_i - \hat{p}_i}{\sqrt{\hat{p}_i(1 - \hat{p}_i)}} \tag{13}$$

.25.2 Influential Observations

Cook's distance and DFBETAS for outlier detection.

.25.3 Goodness of Fit

- Hosmer-Lemeshow test
- Calibration plots
- ROC curves and AUC

.26 Prediction Evaluation Metrics

.26.1 Scoring Rules

Brier Score:

$$BS = \frac{1}{N} \sum_{i=1}^{N} (p_i - y_i)^2 \tag{14}$$

Log Loss:

$$LL = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log p_i + (1 - y_i) \log(1 - p_i)] \tag{15}$$

.26.2 Calibration Metrics

Expected Calibration Error (ECE):

$$ECE = \sum_{m=1}^{M} \frac{|B_m|}{n} |\text{acc}(B_m) - \text{conf}(B_m)| \tag{16}$$

.27 Computational Details

.27.1 Software and Packages

- Python: scikit-learn, statsmodels, lifelines
- R: lme4, survival, rms
- Stan: Bayesian inference
- SQLite: Data storage and querying

.27.2 Convergence Diagnostics

For MCMC chains:

- $\hat{R}$ statistic (Gelman-Rubin)
- Effective sample size
- Trace plots

.27.3 Computational Complexity

[TO COMPLETE: Runtime analysis, memory requirements, scalability]

.28 Bootstrap and Resampling

.28.1 Confidence Intervals

Percentile method for non-parametric CIs:

$$CI_{1-\alpha} = [\hat{\theta}_{\alpha/2}, \hat{\theta}_{1-\alpha/2}] \tag{17}$$

.28.2 Bias Correction

BC_a method for skewed distributions.

.29 Multiple Testing Corrections

Benjamini-Hochberg procedure for FDR control:

$$\alpha_{BH} = \frac{i}{m} \cdot \alpha \tag{18}$$

.30 Power Analysis

[TO COMPLETE: Sample size calculations, effect size detection]

.31 Sensitivity Analysis

[TO COMPLETE: Robustness to model assumptions, prior sensitivity]