Applied Health Analytics and Informatics Using SAS®

Joseph M. Woodside

The correct bibliographic citation for this manual is as follows: Woodside, Joseph M. 2018. *Applied Health Analytics and Informatics Using SAS®*. Cary, NC: SAS Institute Inc.

Applied Health Analytics and Informatics Using SAS®

Copyright © 2018, SAS Institute Inc., Cary, NC, USA

978-1-62960-881-5 (Hardcopy)
978-1-63526-616-0 (Web PDF)
978-1-63526-614-6 (epub)
978-1-63526-615-3 (mobi)

All Rights Reserved. Produced in the United States of America.

For a hard copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government License Rights; Restricted Rights: The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication, or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a), and DFAR 227.7202-4, and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

November 2018

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

SAS software may be provided with certain third-party software, including but not limited to open-source software, which is licensed under its applicable third-party software license agreement. For license information about third-party software distributed with SAS software, refer to **http://support.sas.com/thirdpartylicenses**.

Contents

About this Book ..... **vii**
Acknowledgments ..... **ix**

Chapter 1: Introduction **1**
Introduction 1
Audience Accessibility 2
Learning Approach 3
Experiential Learning Activity: Learning Journal 8

Chapter 2: Health Anamatics **9**
Chapter Summary 9
Chapter Learning Goals 9
Health Anamatics 9
Health Informatics 14
Experiential Learning Activity: Telemedicine 19
Health Analytics 20
Health Anamatics Architecture 26
Experiential Learning Activity: Evidence-Based Practice and Research 28
Health Anamatics Careers 29
Experiential Learning Activity: Health Anamatics Careers 30
Learning Journal Reflection 32

Chapter 3: Sampling Health Data **33**
Chapter Summary 33
Chapter Learning Goals 33
Health Anamatics Process 34
Health Anamatics Tools 34
SEMMA: Sample Process Step 36
SAS OnDemand for Academics Setup 38
Experiential Learning Application: Health and Nutrition Sampling 39
Experiential Learning Application: Health and Nutrition Data Partitioning 55
Experiential Learning Application: Claim Errors Rare-Event Oversampling 62
Learning Journal Reflection 64

Chapter 4: Discovering Health Data Quality **65**
Chapter Summary 65
Chapter Learning Goals 65
Healthcare Quality 66
Experiential Learning Activity: Healthcare Data Quality Check 70

Healthcare Data Quality Case Study 70
Six Sigma Health Data Quality 71
Experiential Learning Activity: Public Data Exploration 72
SEMMA: Exploration 73
Experiential Learning Activity: Health Data Surveillance 75
SEMMA: Modify 76
Experiential Learning Application: Heart Attack Payment Data 77
Experiential Learning Application: Data Quality Exploration 97
Learning Journal Reflection 98

Chapter 5: Modeling Patient Data 101

Chapter Summary 101
Chapter Learning Goals 102
Patients 102
Patient Anamatics 102
Patient Data 102
Healthcare Technology Disruption 104
Experiential Learning Activity: Personal Health Records 106
SEMMA: Model Process Step 107
Experiential Learning Application: Caloric Intake Simple Linear Regression 110
Experiential Learning Application: Caloric Intake Multiple Linear Regression 113
Model Summary 115
Experiential Learning Application: mHealth Heart Rate App 116
Experiential Learning Application: Inpatient Utilization - HCUP 118
Reflection 120

Chapter 6: Modeling Provider Data 123

Chapter Summary 123
Chapter Learning Goals 124
Providers 124
Provider Anamatics 125
Provider Data 126
EHR Implementations 127
EHR Implementation and Success Factors 128
EHR Implementation Process 131
Experiential Learning Activity: Electronic Health Records 133
SEMMA: Model 134
Experiential Learning Application: Hospital-Acquired Conditions 137
Model Summary 155
Experiential Learning Application: Immunizations 155
Learning Journal Reflection 158

Chapter 7: Modeling Payer Data 159
Chapter Summary 159
Chapter Learning Goals 159
Payers 160
Payer Anamatics 161
Payer Data 162
Claim Forms 163
Experiential Learning Activity - Claim Forms Billing 164
Experiential Learning Activity: Claims Adjudication Processing 165
Electronic Data Interchange 167
Experiential Learning Activity: EDI Translation 173
SEMMA: Model 174
Experiential Learning Application: Patient Mortality Indicators 177
Model Summary 197
Experiential Learning Application: Self-Reported General Health 197
Learning Journal Reflection 199
Chapter 8: Modeling Government Data 201
Chapter Summary 201
Chapter Learning Goals 202
Government Agencies 202
Government Health Anamatics 203
Government Regulations 204
Experiential Learning Activity: Government Data Sharing 207
Government Billing and Payments 209
Experiential Learning Activity: Billing Issues and Fraud and Abuse 210
SEMMA: Model 211
Experiential Learning Application: Fraud Detection 215
Model Summary 235
Experiential Learning Application: Hospital Readmissions 235
Learning Journal Reflection 238
Chapter 9: Health Administration and Assessment 241
Chapter Summary 241
Chapter Learning Goals 242
Health Anamatics Administration 242
Code Sets 244
Security 246
Privacy 247
Experiential Learning Activity: HIPAA Administration 249
SEMMA: Assess 251
Experiential Learning Application: Health Risk Score 252
Assess Summary 270

Experiential Learning Application: Hip Fracture Risk 270
Learning Journal Reflection 273
Chapter 10: Modeling Unstructured Health Data 275
Chapter Summary 275
Chapter Learning Goals 275
Unstructured Health Anamatics 276
Social Media 276
Experiential Learning Activity: Social Media Policy 278
Social Media Maturity 279
Experiential Learning Activity: Dr. Google 280
Text Mining 281
Experiential Learning Application: U.S. Presidential Speeches 285
Model Summary 305
Experiential Learning Application: Healthcare Legislation Tweets 306
Learning Journal Reflection 308
Chapter 11: Identifying Future Health Trends and High-Performance Data Mining .. 309
Chapter Summary 309
Chapter Learning Goals 310
Population and Consumer Changes 310
Artificial Intelligence and Robotics Automation 313
Experiential Learning Activity: Robotic Surgery 314
Healthcare Globalization and Government 316
Public Health 318
Big Data Health Anamatics 319
Big Data and High-Performance Data Mining Model 321
Experiential Learning Application: SIDS 323
Model Summary 333
Healthcare Digital Transformation 334
Experiential Learning Application: Lifelogs 336
Learning Journal Reflection 339
Experiential Learning Application: Health Anamatics Project 340
References 343
Index 363

About This Book

What Does This Book Cover?

Health Anamatics is formed from the intersection of data analytics and health informatics. There is significant demand to take advantage of increasing amounts of data by using analytics for insights and decision-making in healthcare. This comprehensive textbook includes data analytics and health informatics concepts along with applied experiential learning exercises and case studies using SAS Enterprise Miner in the healthcare industry setting. The intersection of distinct areas enables connections between data analytics, clinical informatics, and technical software to maximize learning outcomes.

Is This Book for You?

This textbook is intended for professionals, lifelong learners, senior-level undergraduates, and graduate-level students, it can be used for professional development courses, health informatics courses, health analytics courses, and specialized industry track courses.

What Are the Prerequisites for This Book?

An introductory statistics course and an introductory computer applications course are the recommended prerequisites for this book. Topics in an introductory statistics course might include descriptive statistics (frequency, central tendency, and variation) and inferential statistics (sampling, probability, correlation, and experimental design). Topics included in an introductory computer applications course might include computer hardware, productivity software (Microsoft Office, Excel, Word), data access and manipulation, and strategic use of technology.

What Should You Know about the Examples?

Experiential learning activities and applications are included in each chapter so that you can gain hands-on experience with SAS in various healthcare disciplines and in real-world settings. The practical nature of this book helps you to integrate healthcare, analytics, and informatics into health anamatics knowledge, skills, and abilities.

Software Used to Develop the Book's Content

SAS Enterprise Miner 14 is the graphical user interface (GUI) software for data mining and analytics.

Example Code and Data

You can access the example code and data for this book by linking to its author page at https://support.sas.com/woodside.

About the Author

Dr. Joseph M. Woodside is an Assistant Professor of Business Intelligence and Analytics at Stetson University teaching undergraduate, graduate, and executive courses on analytics, health informatics, business analysis, and information systems. He has been a SAS user for over ten years and is responsible for updating the analytics learning goals and course content for the SAS Joint Certificate Program. Before accepting the Business Intelligence and Analytics position at Stetson, Dr. Woodside worked with KePRO, a national healthcare management company, as the Vice President of Health Intelligence, with responsibility for healthcare applications, informatics, business intelligence, data analytics, customer relationship management, employee wellness online platforms, cloud-based systems deployment strategy, technology roadmaps, database management systems, multiple contract sites, and program management. Dr. Woodside previously held positions with Kaiser Permanente, with responsibility for HIPAA Electronic Data Interchange (EDI), national claims and electronic health record implementations, National Provider Identifiers, cost containment financial analytics, and various data analytic initiatives. Learn more about this author by visiting his author page at http://support.sas.com/woodside. There you can download free book excerpts, access example code and data, read the latest reviews, get updates, and more.

We Want to Hear from You

SAS Press books are written *by* SAS Users *for* SAS Users. We welcome your participation in their development and your feedback about SAS Press books that you are using. Please visit sas.com/books to do the following:

- Sign up to review a book.
- Recommend a topic.
- Request information on how to become a SAS Press author.
- Provide feedback on a book.

Do you have questions about a SAS Press book that you are reading? Contact the author through saspress@sas.com or https://support.sas.com/author_feedback.

SAS has many resources to help you find answers and expand your knowledge. If you need additional help, see our list of resources: http://sas.com/books.

Chapter 1: Introduction

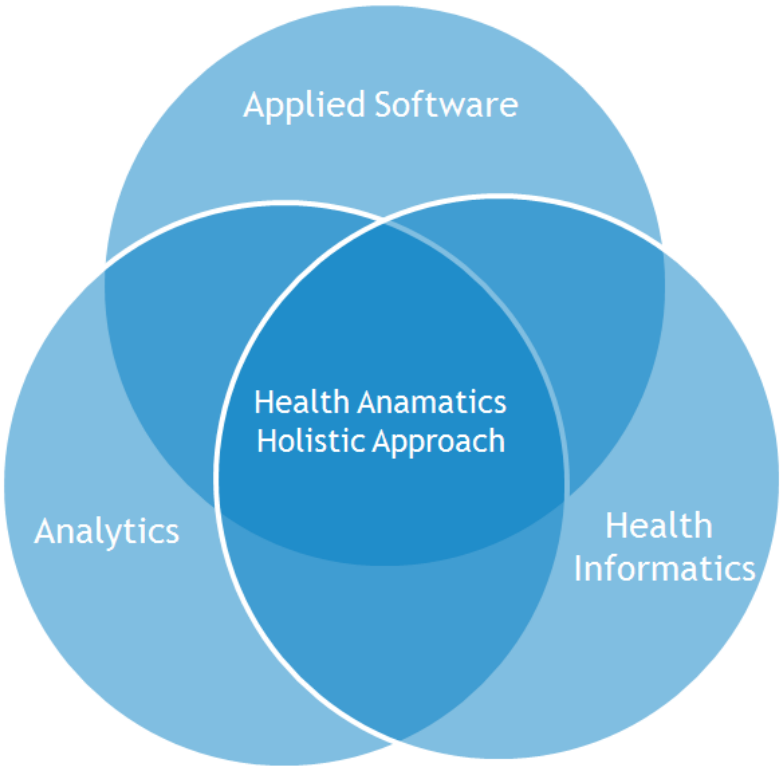
Introduction 1
Audience Accessibility 2
Learning Approach 3
Experiential Learning Activity: Learning Journal 8

Introduction

Health Anamatics is formed from the intersection of data analytics and health informatics. Healthcare systems generate nearly 1/3 of the world's data, and healthcare stakeholders are promised a better world through data analytics and health informatics by eliminating medical errors, reducing re-admissions, providing evidence-based care, demonstrating quality outcomes, and adding cost-efficient care among others. Although healthcare has traditionally lagged behind other industries, the turning point is near with an increased focus across the healthcare sector by way of cost pressures, new technologies, population changes, and government initiatives. There is significant demand to take advantage of increasing amounts of data by using analytics for insights and decision making in healthcare. Healthcare costs keep rising and we can use our technology and analytics capabilities to help address these costs while also improving quality of care. It is our aim to use our knowledge for good and worthwhile causes.

Having conducted several health analytics and informatics related courses and professional education workshops, I have found a need for a comprehensive and current textbook that combines the applied analytics knowledge using SAS with the clinical healthcare informatics concepts. In addition to my ten years of healthcare industry experience, I have met with over 50 industry organizations and executives over the last several years to research relevant content, topics, and applications for health anamatics. This textbook provides a distinguishing feature as a holistic approach as shown in Figure 1.1.

Figure 1.1: Health Anamatics Textbook Distinguishing Approach



Related resources have a primary focus on clinical informatics, technical software, or analytics aspects exclusively, without a connection between all areas to integrate knowledge and maximize learning outcomes.

This textbook contains content and learning objectives, including data analytics and health informatics concepts along with applied experiential learning exercises and case studies using SAS Enterprise Miner within the healthcare industry setting. All clinical data sets are designed to follow the same data structure, data variable set, data characteristics, and methods of published research and industry applied experiential learning examples.

Audience Accessibility

Healthcare and analytics are among the fastest growing areas in industry and curriculum development. This textbook is intended for professionals, lifelong learners, upper-level undergraduates, graduate level students, and can be used for professional development courses, health informatics courses, health analytics courses, and specialized industry track courses. At the graduate level there are currently over 125 analytics programs for which this could be an applied elective or track course, along with over 100 informatics programs for which this could be a core course.

Sample University and Professional Education course titles and current coverage includes:

- Health Anamatics
- Health Informatics
- Health Information and Analytics Management
- Health Analytics
- Healthcare Analytics Management
- Evidence-Based Healthcare Management
- Healthcare Managerial Decision Making
- Applied Analytics in Healthcare

In previous courses, I have had the opportunity to enroll students from a wide variety of specialty areas with a strong interest in learning healthcare and analytics and have helped them be successful in the applied topics. This textbook follows my teaching approach in being accessible to a wide variety of backgrounds and specialty areas including industry professionals, administrators, clinicians, and executives. Examples of major specialty areas from prior enrollment include nursing, information technology, business, international studies, entrepreneurship, sports management, finance, biology, economics, marketing, accounting, and mathematics.

Learning Approach

You might be familiar with the 2015 Disney film, *Inside Out*, which follows the main character Riley, and her emotions of Joy, Sadness, Anger, Disgust, and Fear (Disney, 2017). Watch the following YouTube clip: "Long Term Memory Clip – Inside Out" https://www.youtube.com/watch?v=V9OWEEuviHE

During the film, Joy and Sadness find themselves stuck in endless banks of long-term memory and have trouble finding their way back to headquarters. That is, they do not know the pathway back. Similarly, suppose you are traveling through an endless forest. How do you find your way back? If you walk the path hundreds or thousands of times, you will find it easier each time to find your way back through a clear trail that you have made over time. After a while it will be easy to follow the trail back and find your way home. Human memory is like a nature trail: through frequent retrieval of information that you are creating a pathway, and if you retrieve the information enough, a clear trail forms. Many times along your journey, you might feel that remembering is impossible and you might be like Sadness – this will never happen! Instead, be positive like Joy – with repeated practice and determination that you will find the pathway! Learning takes tremendous effort. It is through this effort that the pathways and memory are built, increasing your intellectual capabilities. Synapses are connected in the brain, and by frequently retrieving memories that you are forming a path to that information. If you retrieve the memory enough times, a well-defined path forms.

Like Riley in *Inside Out*, mental models are psychological representations of real, hypothetical, or imaginary situations, and the individual representation that is used for reasoning. Mental models allow users to understand phenomena, make inferences, respond appropriately to a situation, and define strategies, environment, problems, technology, and tasks. Mental models influence behavior and create reasoning basis, which improve human decision making, by allowing pre-defined models which speed information processing. Mental-model maintenance occurs when new information is incorporated into existing mental models and reinforcement occurs. Mental-model building occurs when mental models are modified based on the new information. Achievement of both mental models is important to achieving

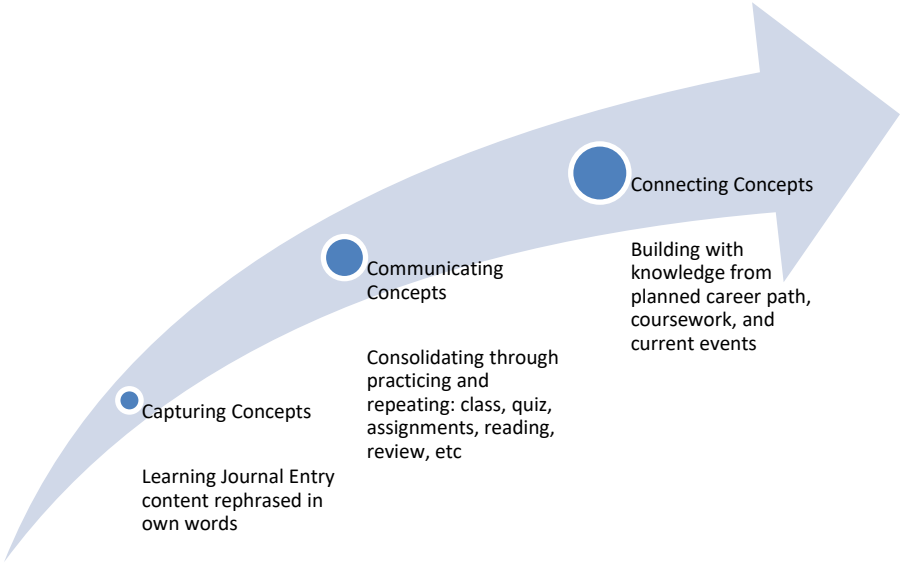
quality and sustained performance. Similarly, health anamatics is intended to provide all stakeholders with high quality, easy to use, and relevant information for decision making. To measure the success, one might gauge whether health anamatics capabilities help users learn. Learning is defined as a purposeful remembering displayed through skillful performance, and is measured as potential change in performance behavior, as the change might occur at a point in time after the information is collected (Vandenbosch and Higgins, 1995; Woodside, 2010a). Health anamatics can be used to improve mental model development. In other words, it help users such as patients, clinicians, and administrators learn.

This textbook follows an experiential, integrative, and applied learning approach using techniques of practice and reflection to reinforce learning. Experiential learning has been included in classrooms as an improved way to educate and engage students as compared with traditional lecture-based learning (Chapman et al., 2016). Traditional education does not offer learners the opportunity to understand the importance of the learning content and real-world scenarios, thereby emphasizing the importance of having learners conduct real-world scenarios to learn and apply to future scenarios. Effective and quality higher education can be achieved only when the balance of academic and practical professional engagement is reached and integrated in a meaningful way (AACSB, 2017). Despite the value of integrative learning across all courses including general education, these student-centered techniques have had limited adoption throughout colleges and universities (Hora, 2017). Instructors and educators also have an important role in experiential learning, requiring individual engagement to facilitate the learning experience and to ensure knowledge generation. Advance planning of the experiential learning activity is critical more so than a traditional lecture, and a learning session might be customized on-the-go and provide opportunities for teachable moments during the session. Experiential learning can also assist with individualized instruction, as each individual has the flexibility to internalize the content to their own individual needs and reflect in a manner meaningful to them as individuals (Roberts, 2003). After following these best learning practices, what can you expect from the results? Research has found that the results include deeper learning and higher grades, which are both agreeable goals.

The empowered and engaged learning approach as shown in Figure 1.2 consists of three phases:

1. Capturing initially difficult concepts through the learning journal and rephrasing in your own words,
2. Communicating concepts through retrieval practice in varying scenarios, this phase is mental model maintenance, and
3. Connecting concepts to professional career areas, industries, and opportunities, this phase is mental-model building where the knowledge is connected to new domains and existing knowledge (Woodside, 2010a).

Figure 1.2: Three-Phase Learning Approach



The learning approach phases can also be thought of as three learning loops, or continuous learning, at each phase. Connections are a key component of experiential, integrative and applied learning that allows connections to be made between concepts and experiences throughout your other courses, professional knowledge, and events, to continually apply your learning to more complex issues and challenges. Over the course of one's career, you will likely change jobs and positions many times. To be successful, you must incorporate your prior knowledge and connect to your new environment to improve decision making and to adapt easily. Initially, the learning might not appear as evident as traditional learning methods would, such as assignments. However, over time the connections become strengthened through experientially based work. The Commission on Accreditation in Physical Therapy Education, the American Association for the Advancement of Science, and the Association of American Colleges and Universities all highlight the critical nature of integrative learning for students to be successful throughout their professional careers. Twenty-first century general education, liberal arts education, co-curricular and pedagogical innovations require effective instructional methods that are able to blend and cut across areas. These methods are the foundation of experiential, integrative and applied learning, and the overall health anamatics approach (Ithaca, 2017; AAC&U, 2017).Trying to solve a problem before being taught the solution leads to better learning, even when errors are made. Applied real-world simulations allow retrieval practices, and spaced and interleaved practice. Interleaved practice often feels slower than massed practice, and as a result, is unpopular and rarely used. Learners might see their grasp of each element coming more slowly and the compensating long-term advantage is not apparent to them. But research shows that mastery and long-term retention is better if you interleave practice, rather than if you mass repeat practice (Brown, et. al, 2014). In this textbook, a common methodology is used in which concepts are interleaved within each chapter. Variable practice is also better, and along with interleaved practice, helps lead to deep learning versus memorization. Reflection is another form of retrieval practice and individual reflection can lead to stronger learning: retrieval knowledge from memory, connecting to new experiences, and visually and mentally rehearsing what you might do differently. Reflection questions might include "What happened?", "What did I do"?, "How did it work out"?, and "What would I do differently next time?". In an effort to assist with learning as you read through this textbook, a summary of learning tips are included below based on best practices (Brown, et. al, 2014; Woodside, 2018a) and shows how this textbook will support those aims:

Table 1.1: Learning Best Practices

Learning Tip	Description	Textbook Alignment
Understand how learning works	Learning consists of three phases: capturing, communicating, and connecting.	At the end of each chapter, there will be a section for capturing your knowledge, practicing communicating your knowledge, and connecting your knowledge.
Spaced repetition and retrieval practice	After reading, ask what are the key ideas, how would I define them, and how do the ideas relate to what I already know. Set aside time every week to quiz yourself on the material. Compared to rereading, this is initially awkward and frustrating when the new learning is difficult to recall week to week. Nevertheless, when you work hard, you strengthen your learning pathways, and spaced practice is more productive than a single session of repeated practice. Spaced repetition and retrieval practice interrupts the process of forgetting and therefore improves learning.	Self-quizzing after reading each section of the textbook. Experiential Learning activities and applications.
Reflection	Reflection is a combination of retrieval and elaboration and generation to add layers to learning and skills. Elaboration is finding additional meaning by explaining concepts in your own words and how these concepts relate to industry events. Generation is trying to solve a problem before seeing the answer, or filling in a blank rather than choosing from a list, and experiential learning activities.	Learning journal entries - Reflect on what you learned during the previous week and how the class learning concepts connects to life outside of the class.
Effort	If learning is easy, then it is quickly forgotten. Some challenges during learning help to make the learning stronger and better remembered, as effortful learning changes the brain, making new connections and increasing intellectual ability. To achieve excellence in learning, you must strive to surpass your current level of ability, striving by its nature results in setbacks that provide the necessary information for learning mastery.	At first, experiential learning might feel difficult. However, through practice and effort, it leads to stronger long-term learning.

Learning Tip	Description	Textbook Alignment
Integrative Learning	According to the Association of American Colleges and Universities (2017), there is an increasing emphasis at a national level to improve student's experiential, integrative, and applied learning through engaged and empowered educational experiences. Integrative learning occurs across disciplines and is critical to deeper learning, instead of localized competencies for a specific field or role. In addition, learning through engagement and student construction of their own ideas is crucial to improve learning, retention, and ability to apply or integrate the knowledge in new areas (AAC&U, 2017). This type of education is one that empowers students to prepare for the wider world and develops transferable intellectual and practical skills, such as analytical and problem solving, with the ability to apply knowledge and skills in real-world settings (AAC&U, 2017).	Experiential and applied learning activities are included within each chapter, and across varying healthcare disciplines and real-world settings to further integrate healthcare, analytics, and informatics into health anamatics knowledge, skills, and abilities.
Rigor	Rigor is a common term that used in higher education for increasing industry preparation of graduates and for increasing academic quality. However, rigor is often defined differently across stakeholder groups and is implemented throughout curriculum in varying ways.	This textbook uses a real-world rigor approach. Real-world rigor is the attainable balance between the possible and impossible, resulting in the constructive conceptualization and realization of student empowerment, engagement and learning through capturing, communicating, and connecting real-world industry knowledge, skills, and abilities to successfully prepare students for their careers and future work environment. Real-world rigor is required to address the adaptability of higher education graduates in a fast-changing business environment and to address the industry competencies that are received in higher education in order to be successful in the future workforce. Changes, such as automation and new skills requirements, will also have major employment impacts in the future (Woodside, 2018a).

Experiential Learning Activity: Learning Journal

Following our learning approach, we will begin with our first learning journal entry. You might record the learning journal entries in an electronic document, a notebook, or a learning management system if available for your course. The learning journal entries will be completed during each class session or as you complete a portion of the textbook. Each learning journal entry should take approximately five minutes. Write as efficiently as possible and continuously for the full time period. You might go back later and edit or add to the learning journal entries as you continue to refresh the topics and build your learning pathways. The learning journal entries will be for your benefit as you proceed through the textbook, as each item is phrased in your own words. The learning journal initially falls into the first phase of capturing concepts. Throughout the textbook and practice, you will begin to consolidate your knowledge through communication, and lastly to connect the concepts through experiential, integrative, and applied learning in order to build your long-term knowledge.

For your first entry, provide your background and knowledge of healthcare, informatics, and analytics. Then rate this knowledge on a scale from 1-100. Lastly, list your goals upon completion of this course or text. For example, this might be your first health-related course and you are seeking to find your area of interest. Or you might have 20 years of experience within a healthcare clinical role and are seeking to expand your knowledge of analytics.

Learning Journal Topics

- Knowledge of Healthcare (1-100)
- Knowledge of Analytics (1-100)
- Knowledge of Informatics (1-100)
- Goals Upon Completion

Ready to take your SAS® and JMP® skills up a notch?

Be among the first to know about new books, special events, and exclusive discounts.
support.sas.com/newbooks

Share your expertise. Write a book with SAS.
support.sas.com/publish

sas.com/books
for additional books and resources.

sas
THE POWER TO KNOW®

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.
Other brand and product names are trademarks of their respective companies. © 2017 SAS Institute Inc. All rights reserved. M1588358 US.0217