

Accelerating Business and Healthcare Decision-Making Processes with Advanced No-Code Data Analytics

Research-in-Progress

David Patrishkoff M.S. LSSMBB – Nova SE University

Robert E. Hoyt FACP FAMIA ABPM-CI - Virginia Commonwealth University

Stephen Bronsburg Ph.D., M.S., M.H.S.A. - Nova SE University

The 2023 Knowledge Management Conference

28 June - 1 July 2023

Hosted by the University of Geneva, Switzerland



No-Code Data Science (ML/AI) is Possible

NO-CODE DATA SCIENCE

Advanced Analytics, Visualizations, Machine Learning, and AI with Open-Source Software



David Patrishkoff
Robert E Hoyt

- No-Code open-source Software (Orange) can apply ML to tabular data, text mining, and image analysis
- Orange predictive analytics capabilities are supplemented by JASP and BlueSky Statistics
- Multi-industry examples and applications
- Free monthly workshops, as we complete our 500-page book this summer
- Certification levels available in the Fall of 2023
- We will link the ChattyPDF chatbot to our book to answer any question related the book content

The Primary Learning Objectives of our No-Code Data Science Textbook

1. Conducting Data preparation and wrangling
2. Constructing compelling data visualizations
3. Developing and evaluating predictive models
4. Conducting time series forecasts and survival analysis
5. Conducting Geolocation-based analysis
6. Exploring the future of Lean Six Sigma Methodologies
7. Implementing Image Analytics Techniques
8. Mastering Text Mining Strategies

We are offering proficiency-based training and certification in no-code data science, equal to the depth, scope, and quality of any data science coding curriculum

The Need for More Data Scientists

Unleash the Data and its Potential with more NCDS Data Scientists and more Progress

Here are some of the reasons for the shortage of Data Scientists:

- Lack of high-end technical and non-technical skilled Data Scientists in the market
- Data Science skills have a steep learning curve with coding and programming being one of the biggest roadblocks for entry in this field
- Data Science programs are often only offered to graduate programs which limits entry for a broader pool of talent into this industry
- Many large companies compete with each other to implement ML/AI which consumes most of the available Data Scientists

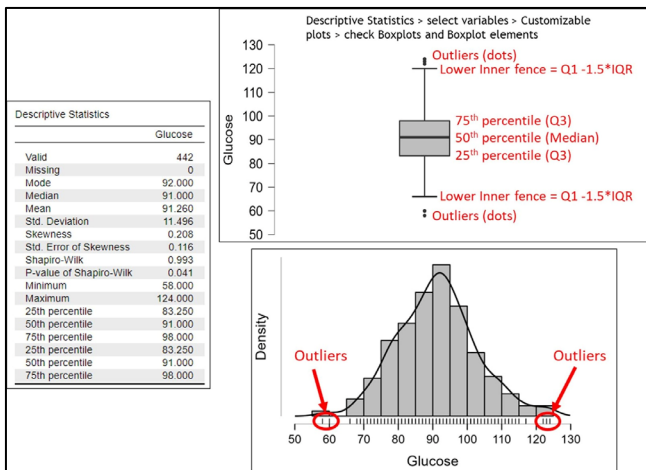
Dilemma:

- Most small and mid-sized organizations cannot afford Data Scientists to help them with ML/AI implementation

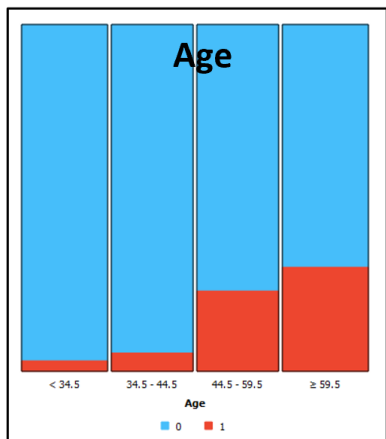
Adding new trained and certified **No-Code Data Scientists** to the market could help to close the skills gap and avoid the coding roadblock in the learning process

1. Data Preparation and Wrangling

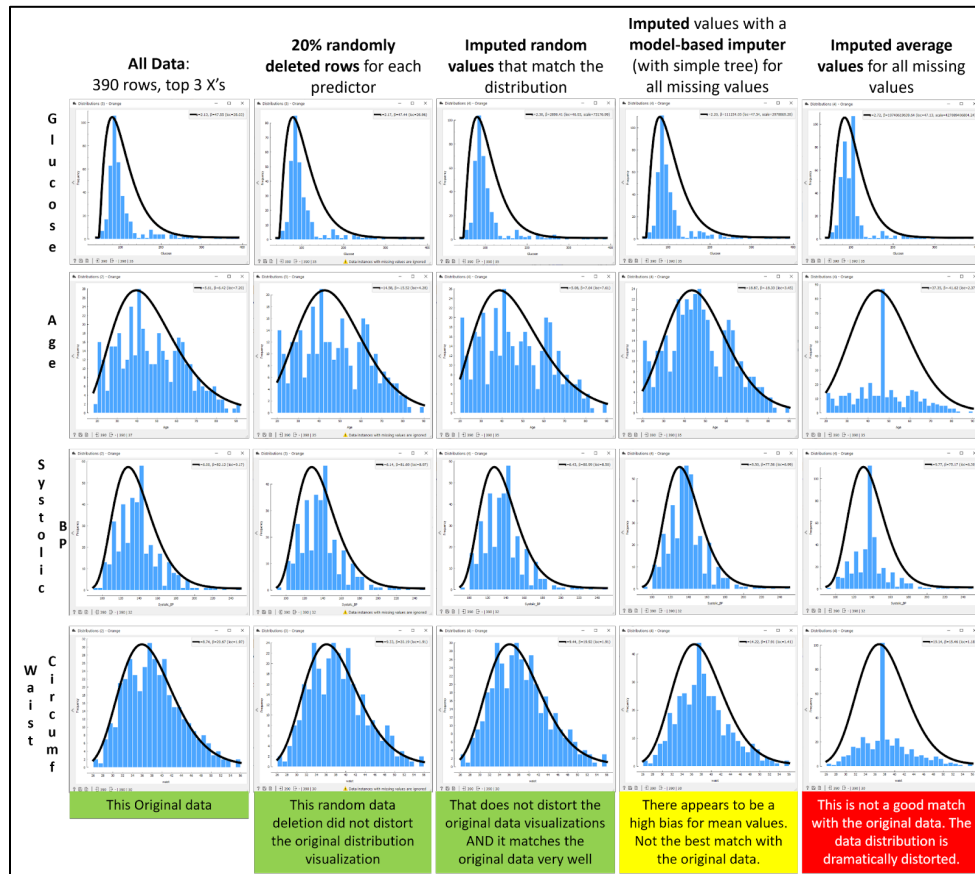
Reshaping Chaos into Clarity: Polishing and Prepping Data for Predictive Modeling



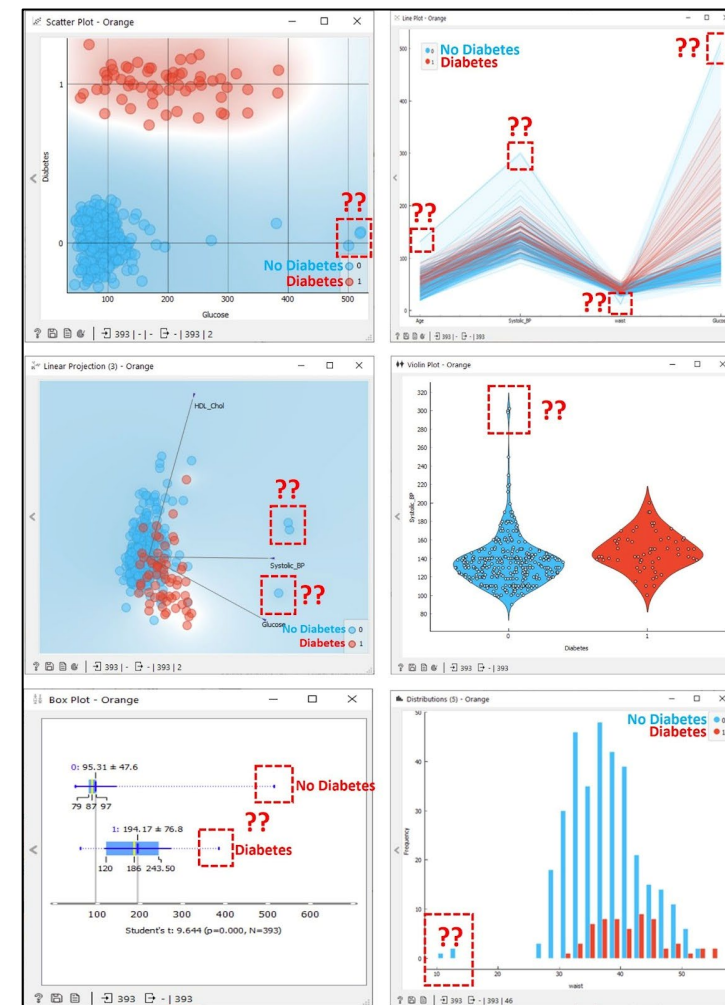
JASP Outlier analysis



Orange data binning



Orange Missing Data Imputation experiments

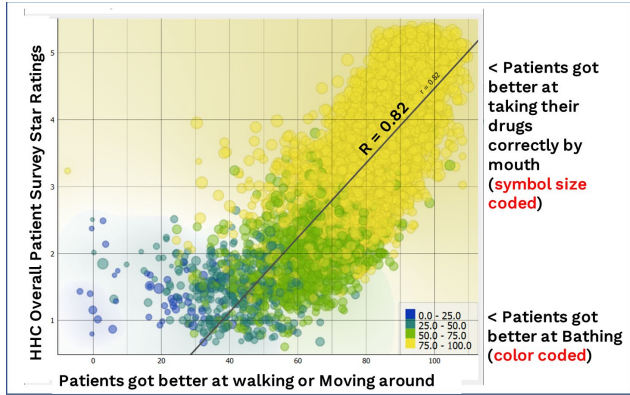


Orange data entry error identification

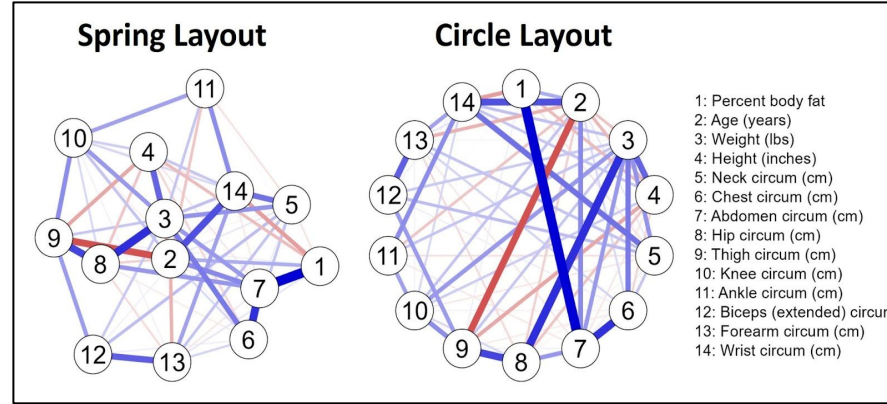
Orange, BlueSky, and JASP offers a wide variety of data prep and wrangling techniques

2. Constructing Compelling Data Visualizations

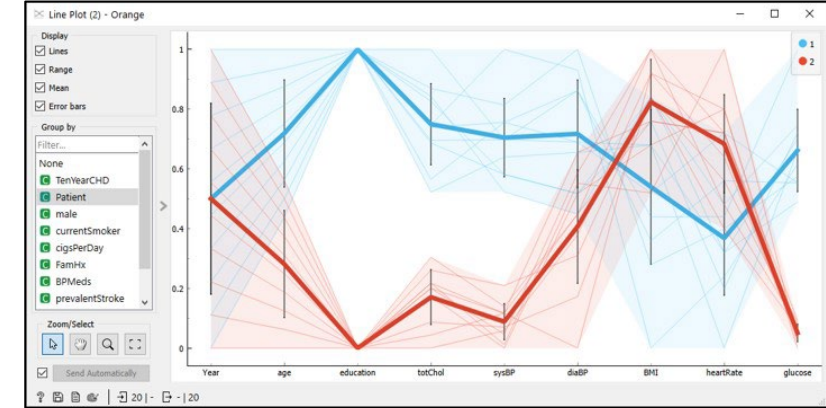
Painting Stories with Numbers



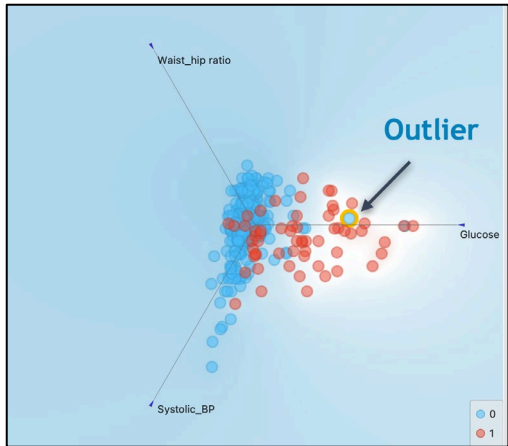
Orange Scatter Plot



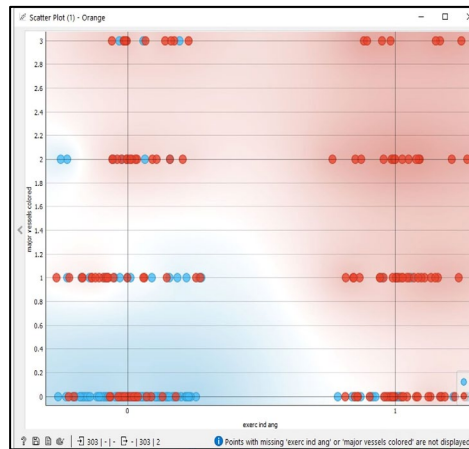
JASP Network / Correlation plots



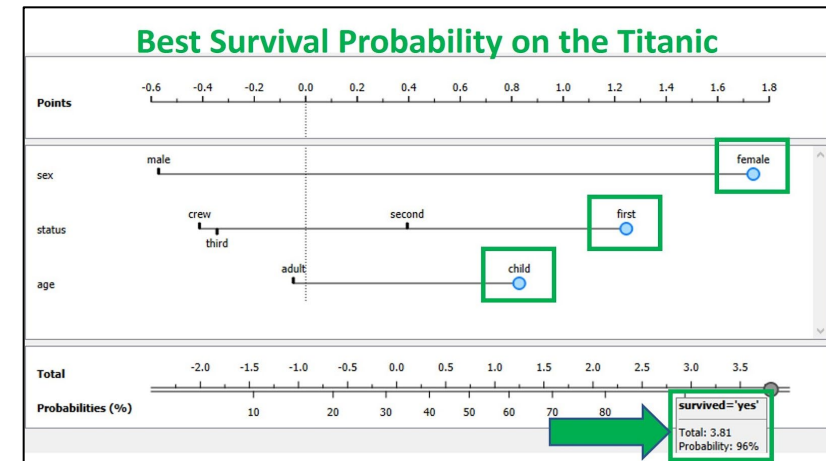
Orange Parallel Coordinates Plot



Orange 3D plot



Orange Scatter Plot

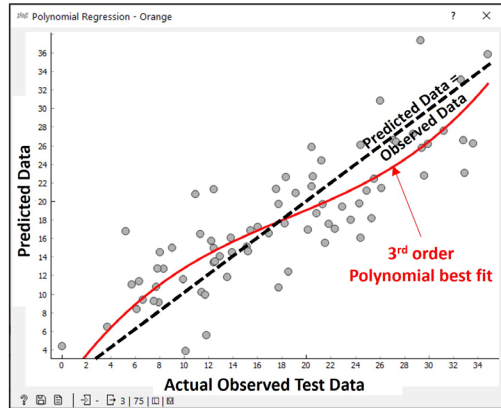


Orange Nomogram

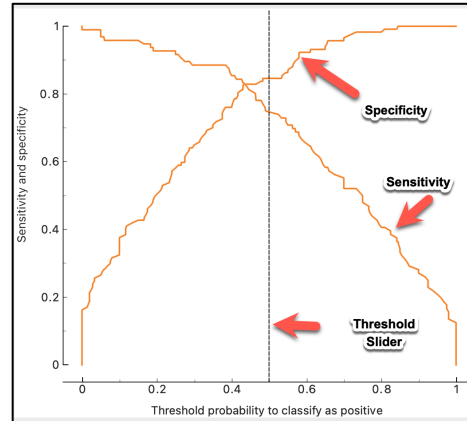
Orange, JASP, and BlueSky Statistics offers a wide variety of data visualizations

3. Developing and Evaluating Predictive Models

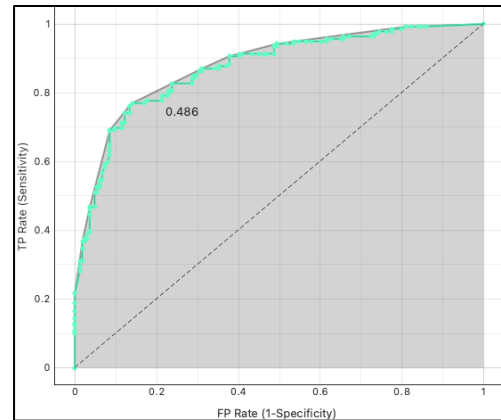
Forging Future Insights: Crafting and Calibrating Predictive Models



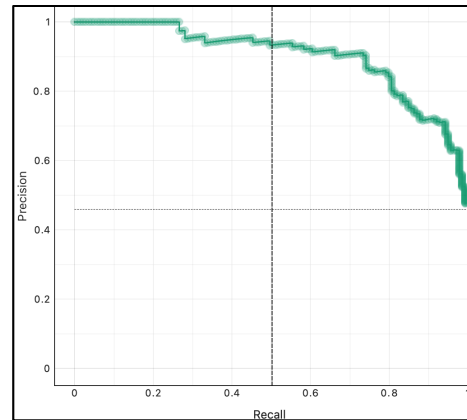
Orange Model Residuals analysis



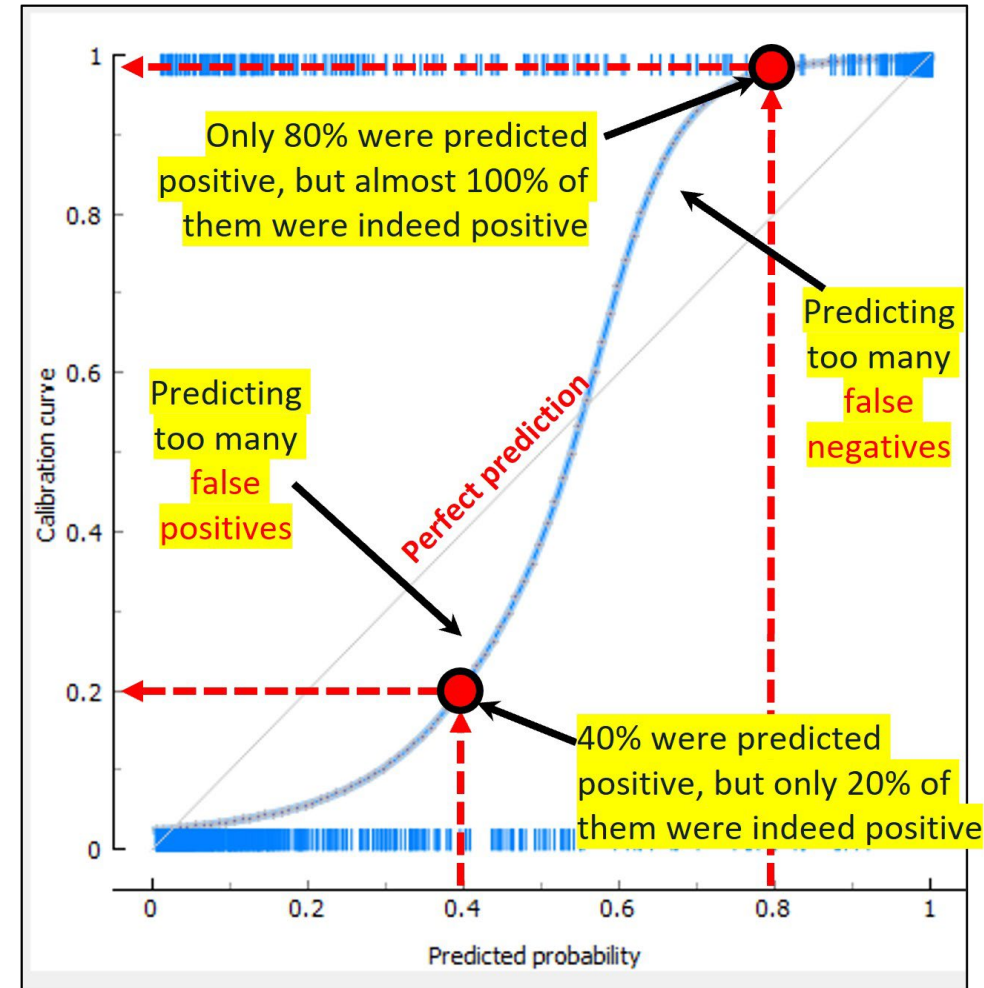
Orange Sensitivity vs Specificity plot with Threshold slider



Orange ROC Curve



Orange Precision-recall Curve

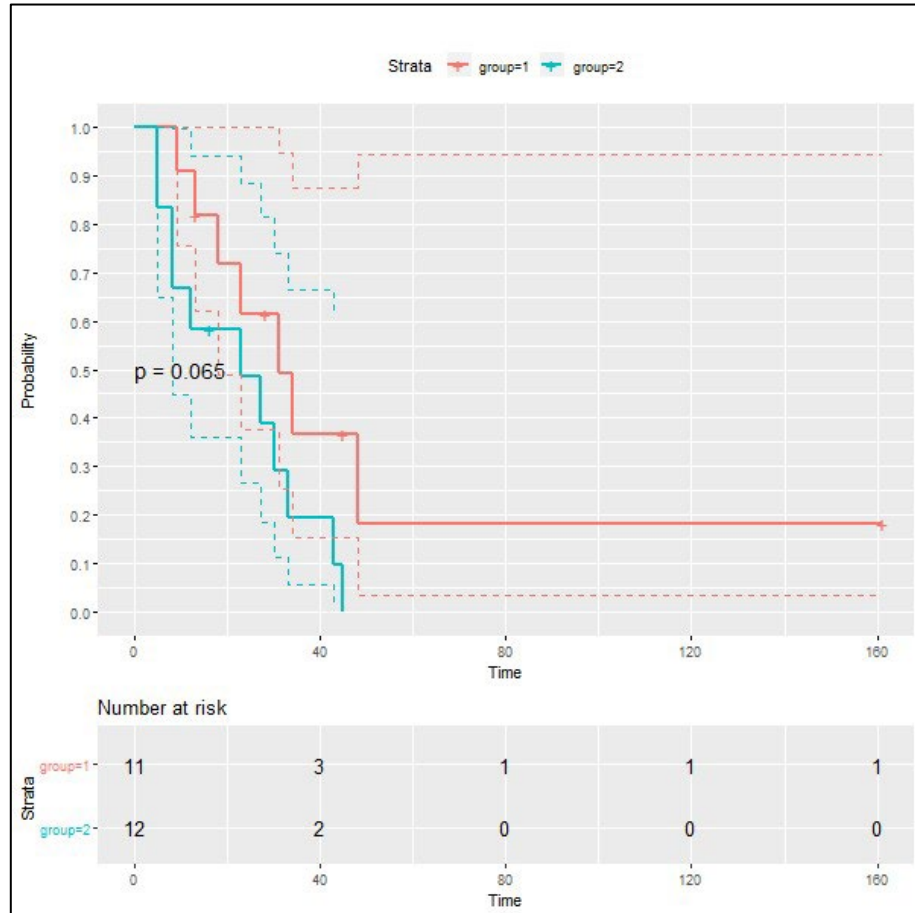


Orange model Calibration Curve

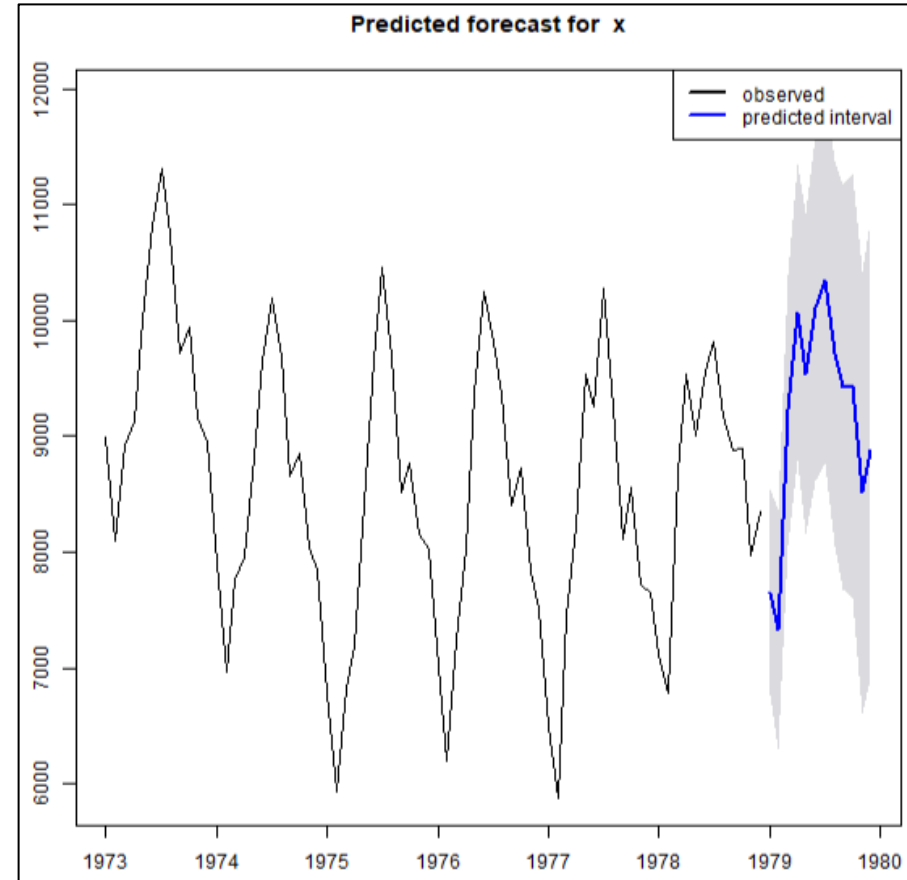
Orange, BlueSky, and JASP offers a wide range of model building and evaluation methods

4. Conducting Time Series Forecast and Survival Analysis

Unraveling Time's Tapestry: Mastering Forecasts and Lifelines with Analysis



Kaplan-Meier Survival estimates in BlueSky

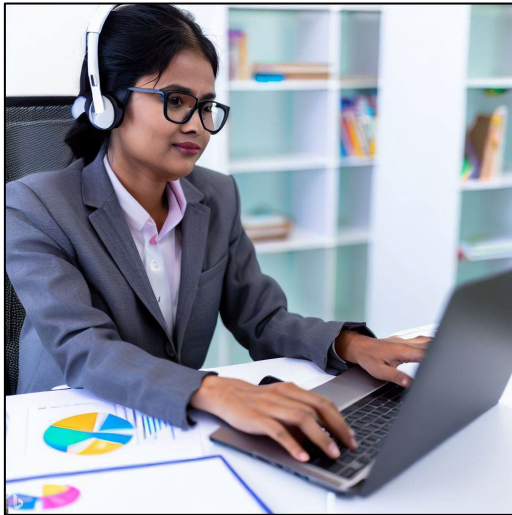


USA Accidental deaths with a 12-month prediction in BlueSky

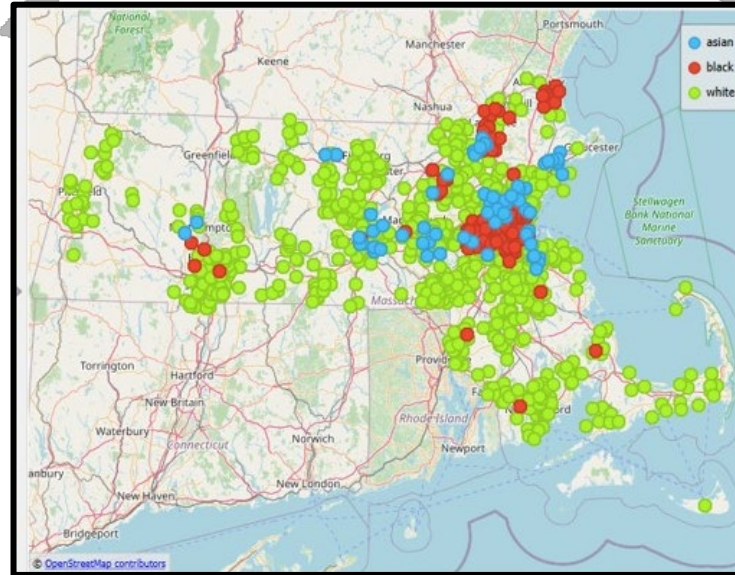
Orange and BlueSky offers Time Series Forecasting and Survival Analysis techniques

5. Conducting Geolocation-based Analysis

Merging Art, Geography, and Analytics



Analytical



Geospatial
Visualizations



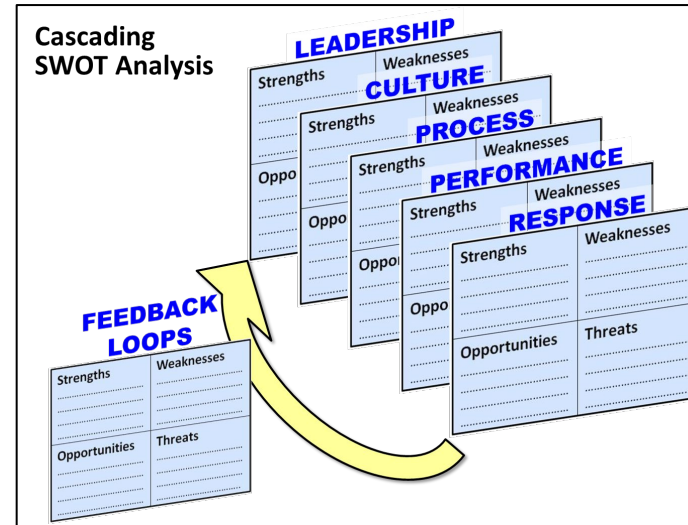
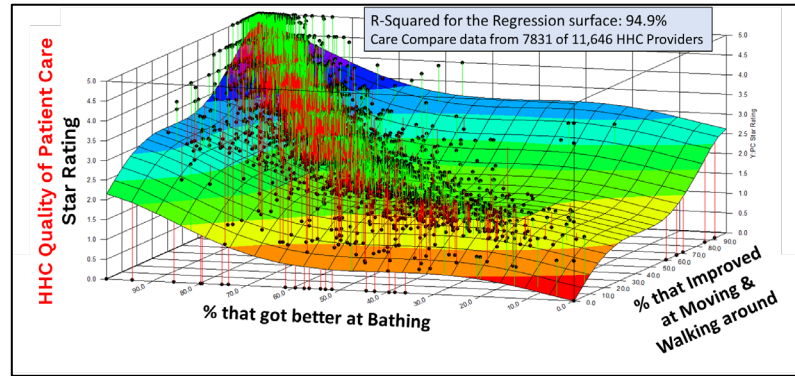
Creative

AI Images created by DALL-E

Orange and BlueSky offers Geo-Spatial Analysis techniques

6. Exploring the Future of Lean Six Sigma (LSS) Methodologies

Revolutionizing Efficiency: Disruptively Reinventing Lean Six Sigma with ML/AI



The future of LSS will focus on:

- Realtime and automated data analysis at the Gemba
- Expanded cascading root cause analysis
- Rapid continuous improvement
- Organizational cultural risk identification and mitigation
- AI chatbot support for all problem brainstorming activities
- Integrated Predictive Analytics
- ChatPDF access to improved processes and procedures

LSS Brainstorming to identify Projects to Reduce Manufacturing Waste

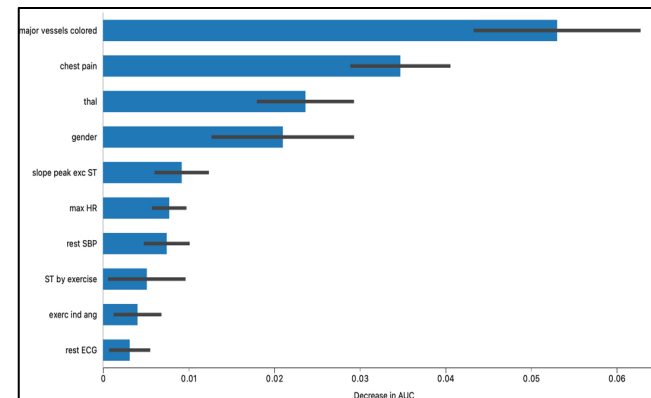
- 1. Defects, scrap, repair, reworking, fixing, mistakes, other errors in a process, missing or incorrect info, Internal / external customer complaints or missed deadlines**
- 2. Non-optimal Processes and NVA:** work, steps or resources used to perform a task that a customer sees no value in. Low capacity or capability, poor communications, lack of training / cross-training, too much done in series, not in parallel, redo loops are present, staff shortages, lack of innovation
- 3. Over-processing or over-production:** doing more than the customer requests, maybe you do not know what they really want, over killing specs, too many redundant quality checks required that a customer would not directly want to pay for.
- 4. NVA Transportation:** movement or loops (actual or virtual) of material, products, inventory, products, paperwork or information
- 5. High Inventories of raw material, VIP and / or finished goods, safety stock Kanban not widely used; JIT delivery not implemented.**
- 6. Waiting:** slow upstream process, bottleneck, bureaucracy, equipment, machine or downtime, changeover times, redo loops, too many adjustments, material not available, actions not done, non-synchronized hand offs, confusing and complex processes, non-motivated staff, high maintenance processes, overburdened staff, can't get a win-win situation, slow development times, non-balanced cycle times, decisions not made, too long between the end of any process and the start of a new process step start

Expanded Root Cause Brainstorming

A fishbone diagram with 'WASTE' at the head. The main spine has several arrows pointing to it from both sides, representing various causes. The causes are numbered 7 through 12.

- 7. Material Flow inefficiencies** along the main flow of process steps for products being manufactured
- 8. Non-competitive Supply Chain:** High cost, low capability and too many suppliers with ineffective audits in place, not challenging & / or monitoring suppliers in their pursuits of effective productivity improvement & cost reduction initiatives
- 9. Inefficient Purchasing methods:** Not using economics-of-scale, no back-up plans, use of cheaper substitute materials with the same quality, consolidated purchases, e-business and e-bidding, etc
- 10. Needless Motion of people,** vehicles, equipment that should be observed and spaghetti charted noting the distance traveled for each step and the total distance traveled.
- 11. New Technology not applied to:** achieve error-proofing, high first time yields, process parameter data, automatic defect tracking for root cause analysis purposes, get higher output manufacturing processes, achieve cost reductions, quality checks
- 12. Variable and/or Fixed Costs are too high for:** labor, rent, borrowing interest rates, material, utilities (water, electric power, steam, etc.), contracted services, raw material, depreciation, benefits, group medical plans, travel expenses, etc.

© Copyright 2008 David Patrishkoff

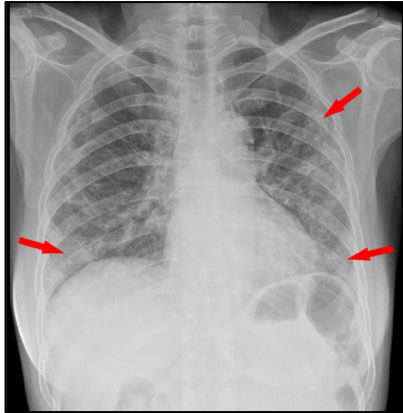


Orange Feature importance results for a classification problem

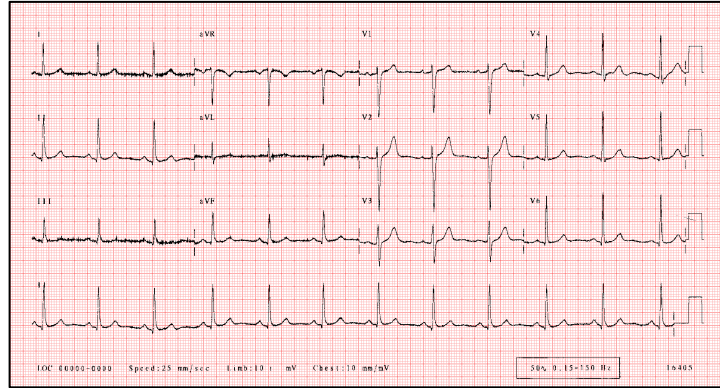
BlueSky covers every standard and advanced Six Sigma Data Analysis Technique

7. Implementing Image Analytics Techniques

Picturing the Future: Predictive Insights with Image Analytics



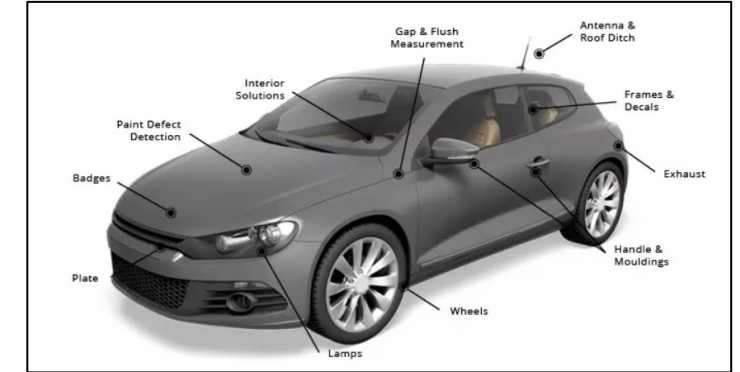
Xray diagnoses



ECG diagnosis



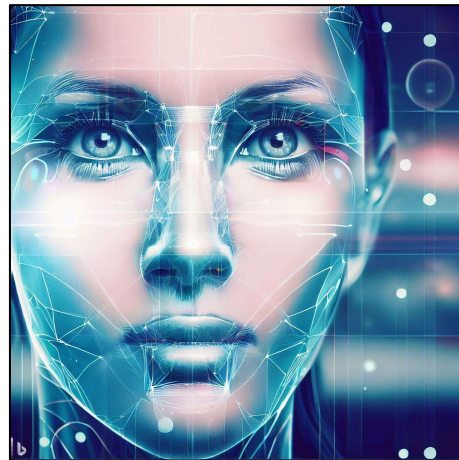
Image detection



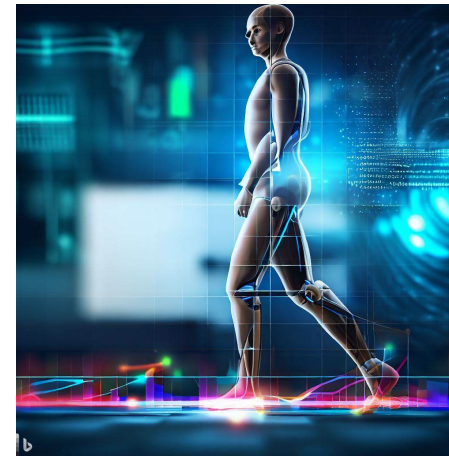
Vehicle quality checks



Mood detection



Facial recognition



Posture classification



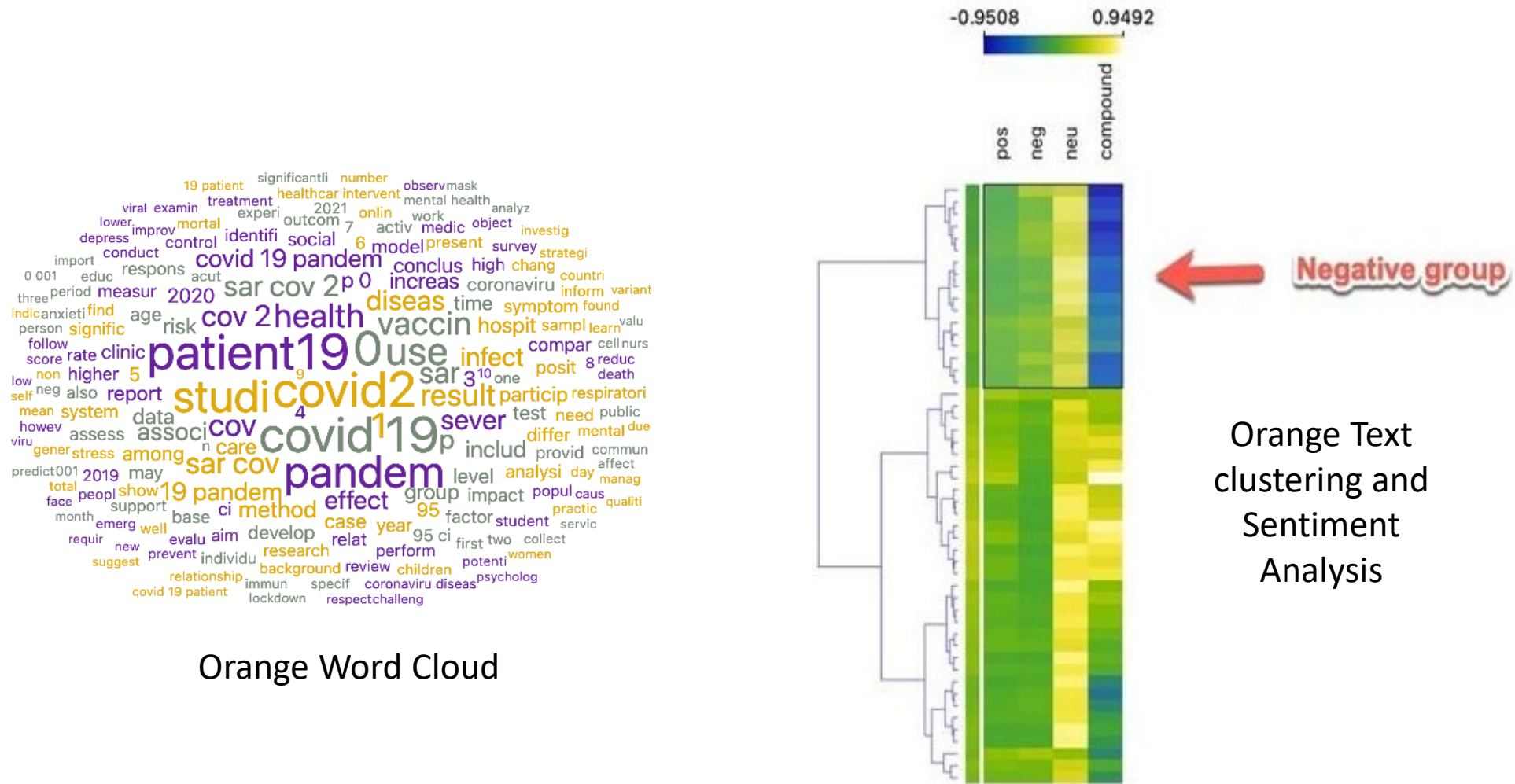
Vehicle classification

All Images created by DALL-E

Orange can provide a wide range of image detection, classification, and predictive analytics

8. Mastering Text Mining Strategies

From Words to Wisdom: Master the Art of Text Mining for Deep Insights



Orange offers a variety of Text mining tools ranging from simple Word Clouds to Text Clustering, Sentiment Analysis, and Text Grouping Predictions

No-Code Data Science Book - Author Bios

David Patrishkoff, M.S. is a Lean Six Sigma Master Black Belt with C-level worldwide executive experiences with engineering, quality, and manufacturing responsibilities at multi-billion-dollar revenue companies. In 2001, he founded a consulting and training company, E3 (www.e3.business), where he has trained and consulted for organizations in over 60 different industries worldwide to resolve their mission-critical issues with innovative process improvement and data analysis techniques. In 2018, he added machine learning techniques to gain more insights into his research of large datasets, which included the Fatality Reporting Analysis System (FARS) for motor vehicle accidents in the USA and the Home Health Outcome Assessment Information Set (OASIS) data for home healthcare patients in the USA. He speaks regularly at international conferences about his research and other topics of interest. He has a contributing faculty position at the Kettering University School of Management in Flint, Michigan. He is also an Adjunct Professor at the Dr. Kiran C. Patel Osteopathic School of Medicine in Ft. Lauderdale, Florida, part of Nova SE University. He has already incorporated Orange into some of the analytics classes he teaches at Nova University. He has trained, certified, and mentored many professionals in various topics, including over 3,000 professionals in Lean Six Sigma techniques. He and his company have also trained over 23,000 healthcare professionals in High-Reliability Organizations (HRO) plans to reduce medical errors in healthcare systems.

Dr. Robert Hoyt FACP FAMIA ABPM-CI is an internal medicine physician who has taught health informatics for the past two decades. He is the editor and author of *Health Informatics: Practical Guide*, which is in its eighth edition. In 2014 he became board certified in clinical informatics; in 2016, he became a Fellow of the American Medical Informatics Association (FAMIA). He has extensive experience in data science and is the editor and author of *Introduction to Biomedical Data Science* (2019) and *Data Preparation and Exploration* (2020). Information about those textbooks can be found at <https://www.informaticseducation.org>.

Dr. Hoyt is the honorary president of the [Medical Intelligence Society](#), contributing to the Data Science Tip of the Month. He is on the Board and is faculty for the [American Board of Artificial Intelligence in Medicine Board Review Course](#). Furthermore, he is a reviewer for multiple medical journals, including the new [Intelligence Based Medicine](#) journal.