



# Future Proof Your Career: What every executive needs to know about Adaptive Intelligence

MAY 16 & 17, 2018

CLEVELAND PUBLIC AUDITORIUM, CLEVELAND, OHIO

[WWW.NEOOUG.ORG/GLOC](http://WWW.NEOOUG.ORG/GLOC)

# **Future Proof Your Career: What every executive needs to know about Adaptive Intelligence**

Great Lakes Oracle Conference 2018

Tim VlamiS

[tvlamis@vlamis.com](mailto:tvlamis@vlamis.com)

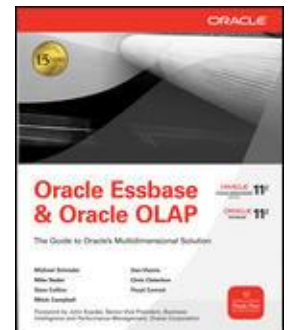
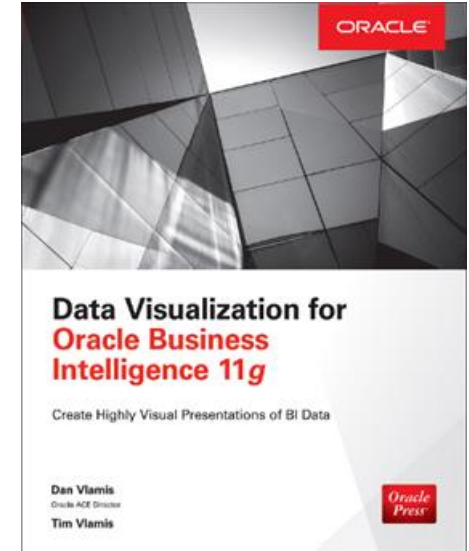
# VlamiS Software Solutions

- VlamiS Software founded in 1992 in Kansas City, Missouri
- Developed 200+ Oracle BI and analytics systems
- Specializes in Oracle-based:
  - Enterprise Business Intelligence & Analytics
  - Analytic Warehousing
  - Data Mining and Predictive Analytics
  - Data Visualization
- Multiple Oracle ACEs, consultants average 15+ years
- [www.vlamiS.com](http://www.vlamiS.com) (blog, papers, newsletters, services)
- Creators of the [Force Directed Graph Plugin](#) on the [Oracle Analytics Library](#)
- Co-authors of book “Data Visualization for OBI 11g”
- Co-author of book “Oracle Essbase & Oracle OLAP”
- Oracle University Partner
- Oracle Gold Partner

 EDUCATION RESELLER


 Gold  
Partner

Specialized  
Oracle Business Intelligence  
Foundation Suite 11g





## Vice President & Analytics Strategist

- 30+ years in business modeling and valuation, forecasting, and scenario analyses
- Oracle ACE  Oracle ACE
- Instructor for Oracle University's Predictive Analytics, Data Mining Techniques, and Oracle R Enterprise Essentials Courses
- Professional Certified Marketer (PCM) from AMA
- MBA Kellogg School of Management (Northwestern University)
- BA Economics Yale University



# Presentation Agenda

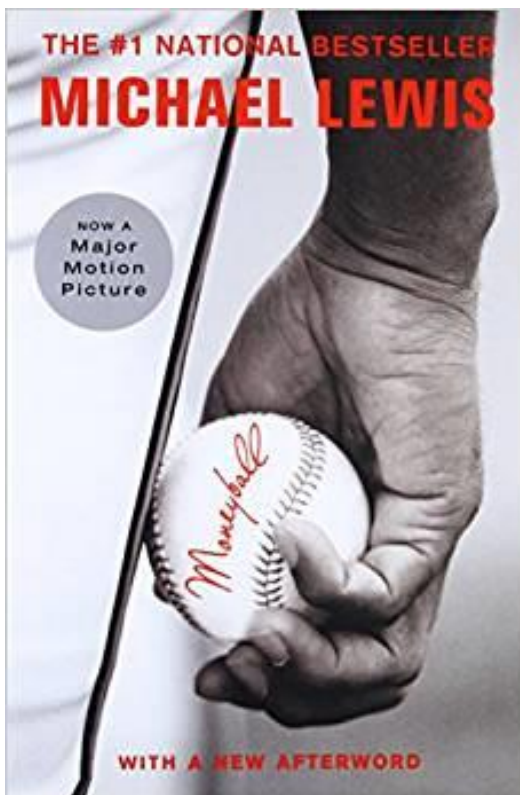
- A little language and background
- Machine Learning Systems
- Basics of Bayesian Thinking
- Sources for learning more
- **YOUR QUESTIONS AT ALL TIMES!**





# Moneyball

- “One absolutely cannot tell, by watching, the difference between a .300 hitter and a .275 hitter. The difference is one hit every two weeks.”





# Many Words, Closely Related

- Adaptive Intelligence → like machine learning with broader inputs
- Machine Learning → like data mining with auto-updated predictions

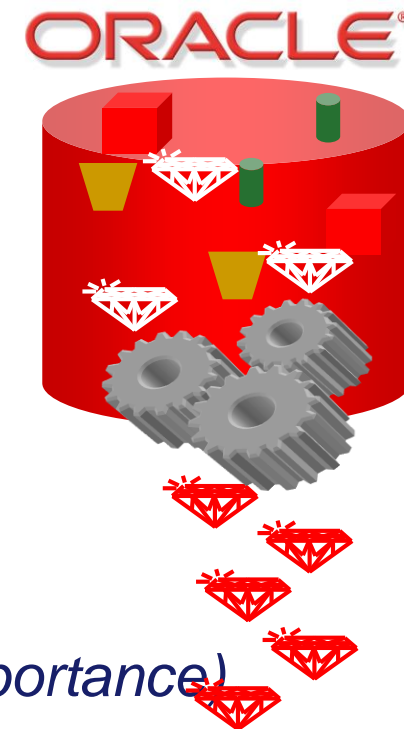
- Data
- Predictive Analytics
- Regression
- Data Mining
- Processing
- Machine Learning

SQL Analytics   Anomaly Detection  
Python   Adaptive Intelligence   Data Science  
Diagnostic Analytics   AI   Advanced Analytics  
Classification  
SQL   R   Clustering   Algorithm   Descriptive Analytics  
Prescriptive Analytics   Artificial Intelligence  
Machine Learning



# What is Machine Learning?

- Automatically sifts through data to find hidden patterns, discover new insights, and make predictions
- Predictions are updated with new information from the system
- Machine Learning algorithms can provide valuable results:
  - Predict customer behavior (*Classification*)
  - Predict or estimate a value (*Regression*)
  - Segment a population (*Clustering*)
  - Identify factors more associated with a business problem (*Attribute Importance*)
  - Find profiles of targeted people or items (*Decision Trees*)
  - Determine co-occurrences and “market baskets” within an event set (*Associations*)
  - Find fraudulent or “rare events” (*Anomaly Detection*)







# Four Realms of Analytics

Probability Based

**Diagnostic  
Analytics**

**Predictive  
Analytics**

Rules Based

**Descriptive  
Analytics**

**Prescriptive  
Analytics**

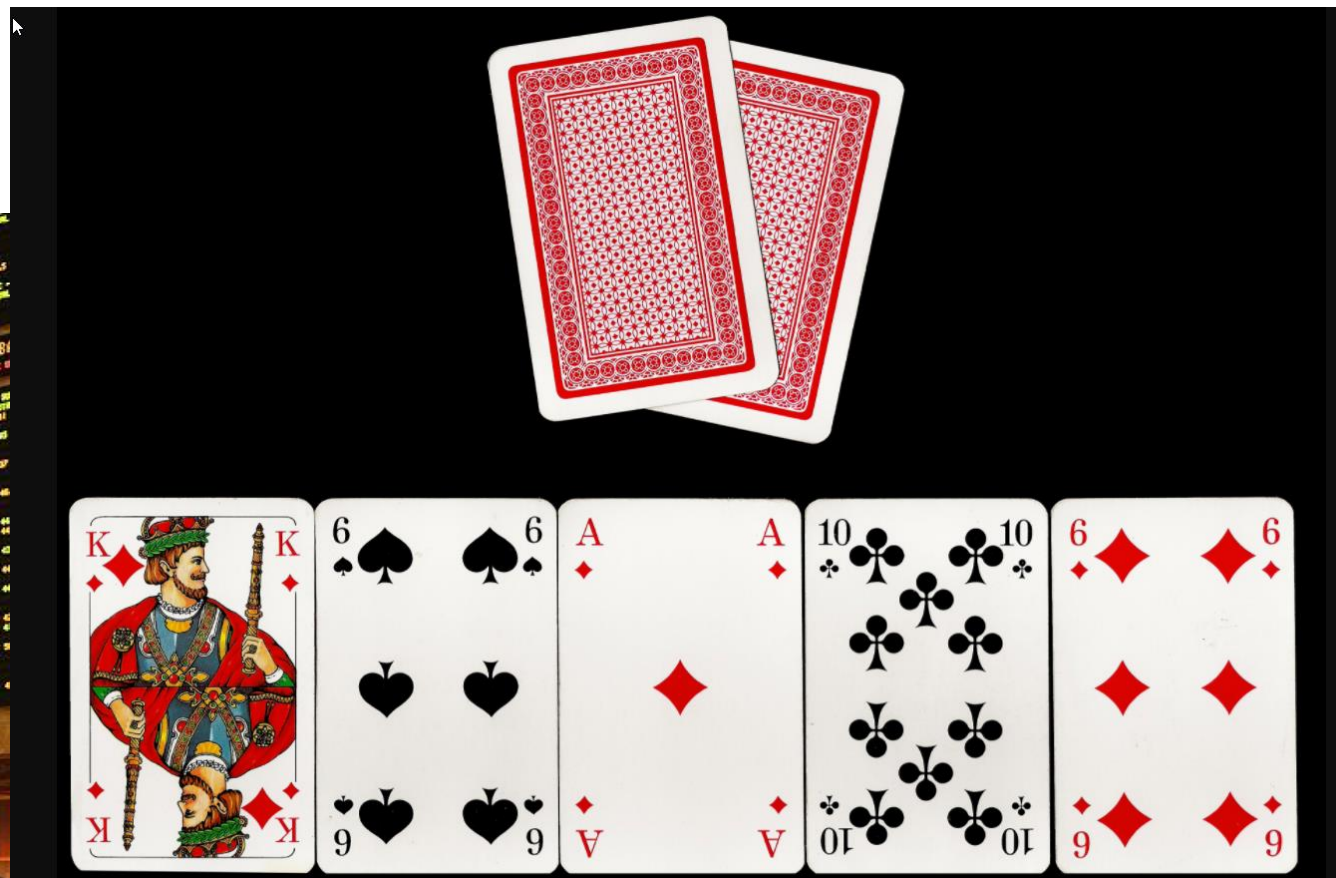
Past

Future



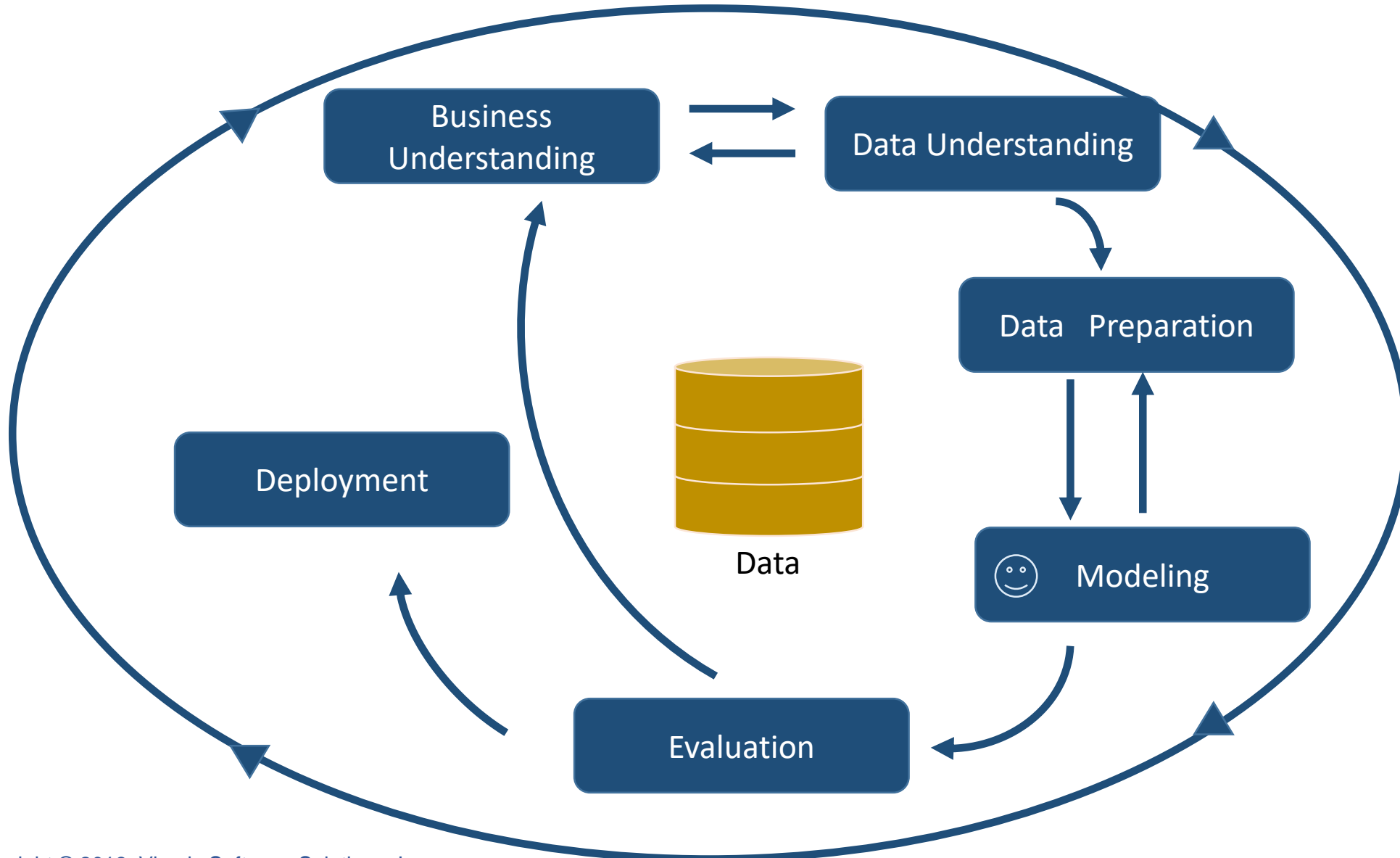
# Probabilities, not Outcomes

- The likelihood of an event defines what is a good bet, not





# CRISP-DM Phases



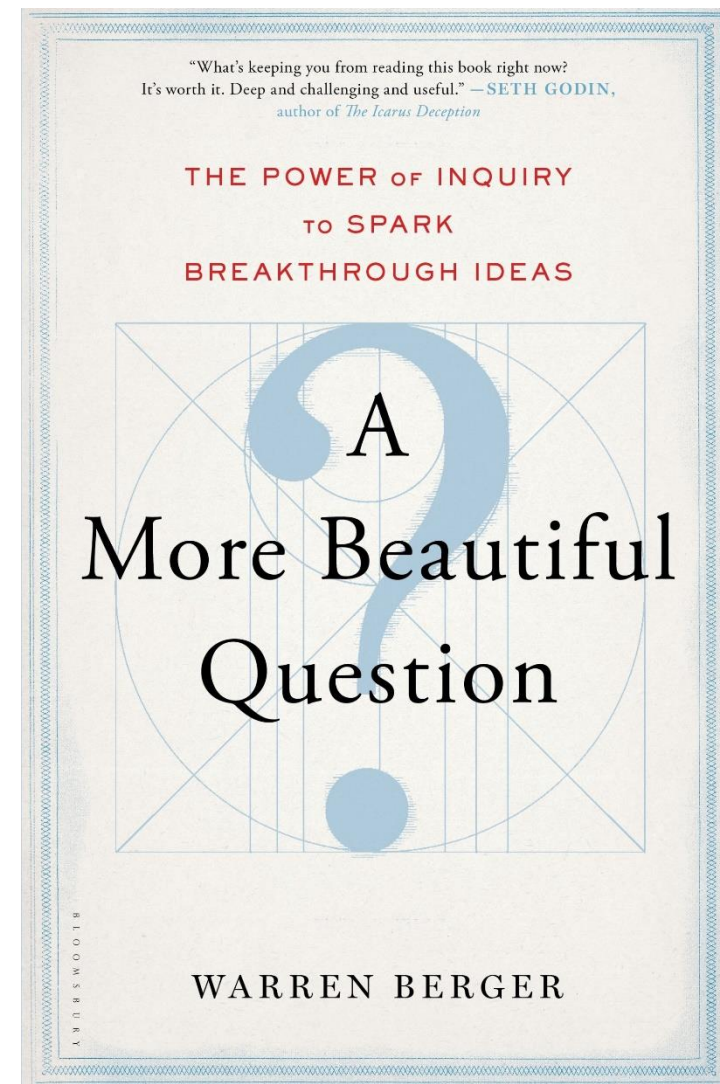


# Good Questions/Hypotheses are Needed

What behaviors in the past year are most significant in terms of segmenting our customers?

What's the Life Time Value of each customer?  
What's a potential new customer worth?

Which products are purchased together most often? Which products are purchased with our most profitable products?







# Listen to Data

- Relative importance
- Natural relationships
- Similarities/differences
- Predictions







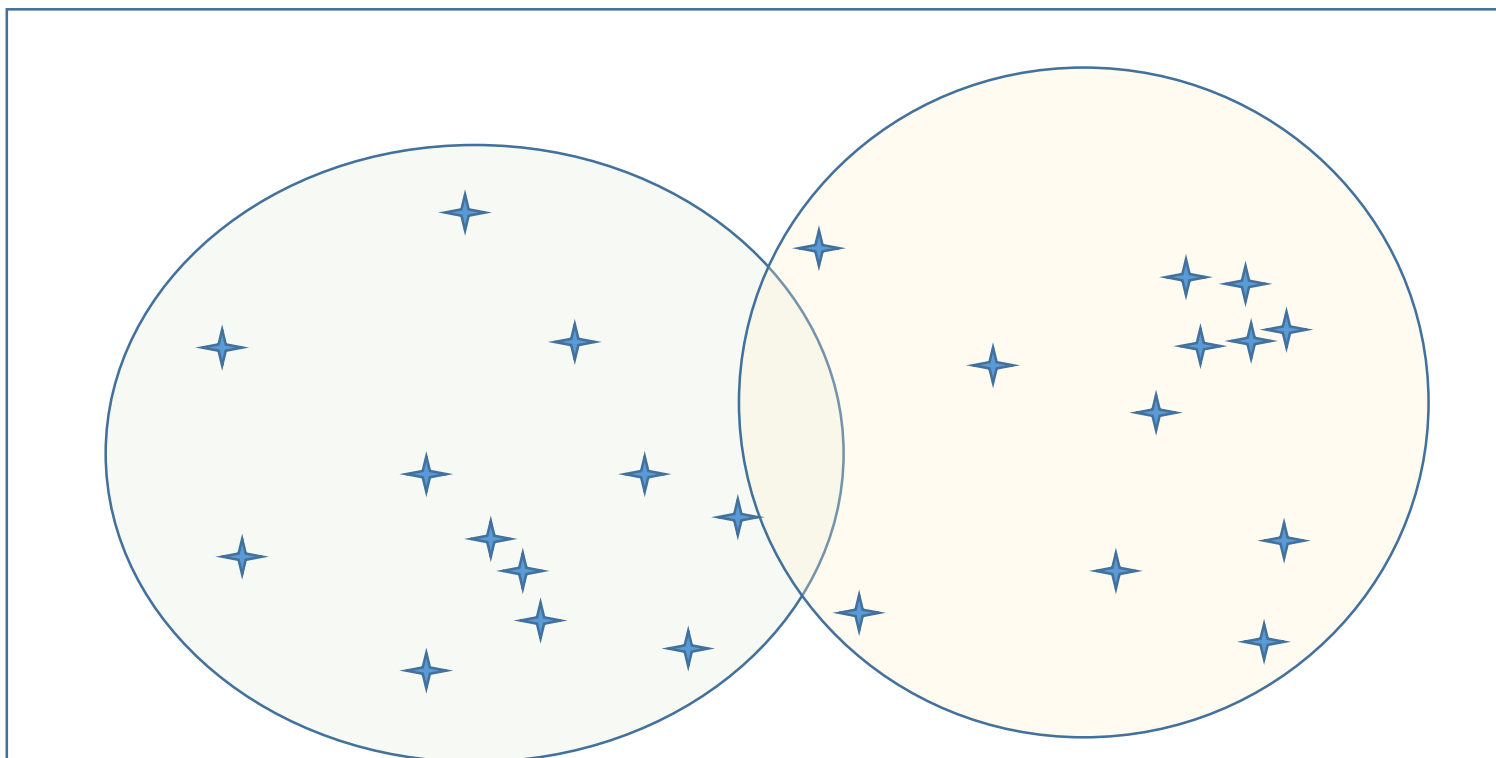
- **Ensure the right data is being applied to the right business problem.**
- **Make sure data is of the best quality, and is double-checked.**
- **Test diligently.**
- **Gradually introduce features helpful to end-users.**

- Source Forbes Magazine's Joe McKendrick interview with Oracle's Jack Berkowitz <https://www.forbes.com/sites/joemckendrick/2018/04/11/4-ways-to-ensure-ai-credibility-and-ensure-adoption/#6037e74e1695>



# What is Clustering?

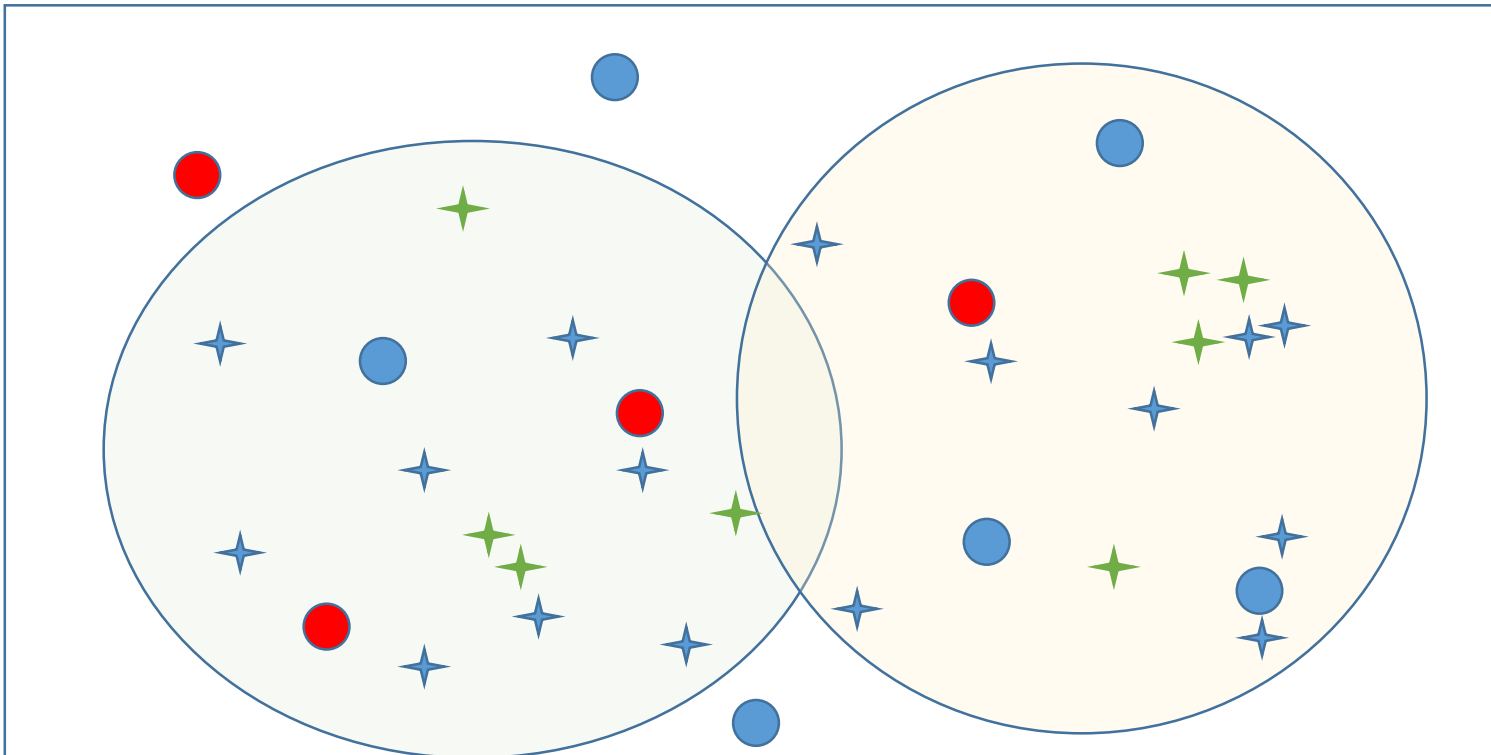
- Dividing a large set into smaller groups of similar and dissimilar members





# More Dimensions Makes Clustering Harder

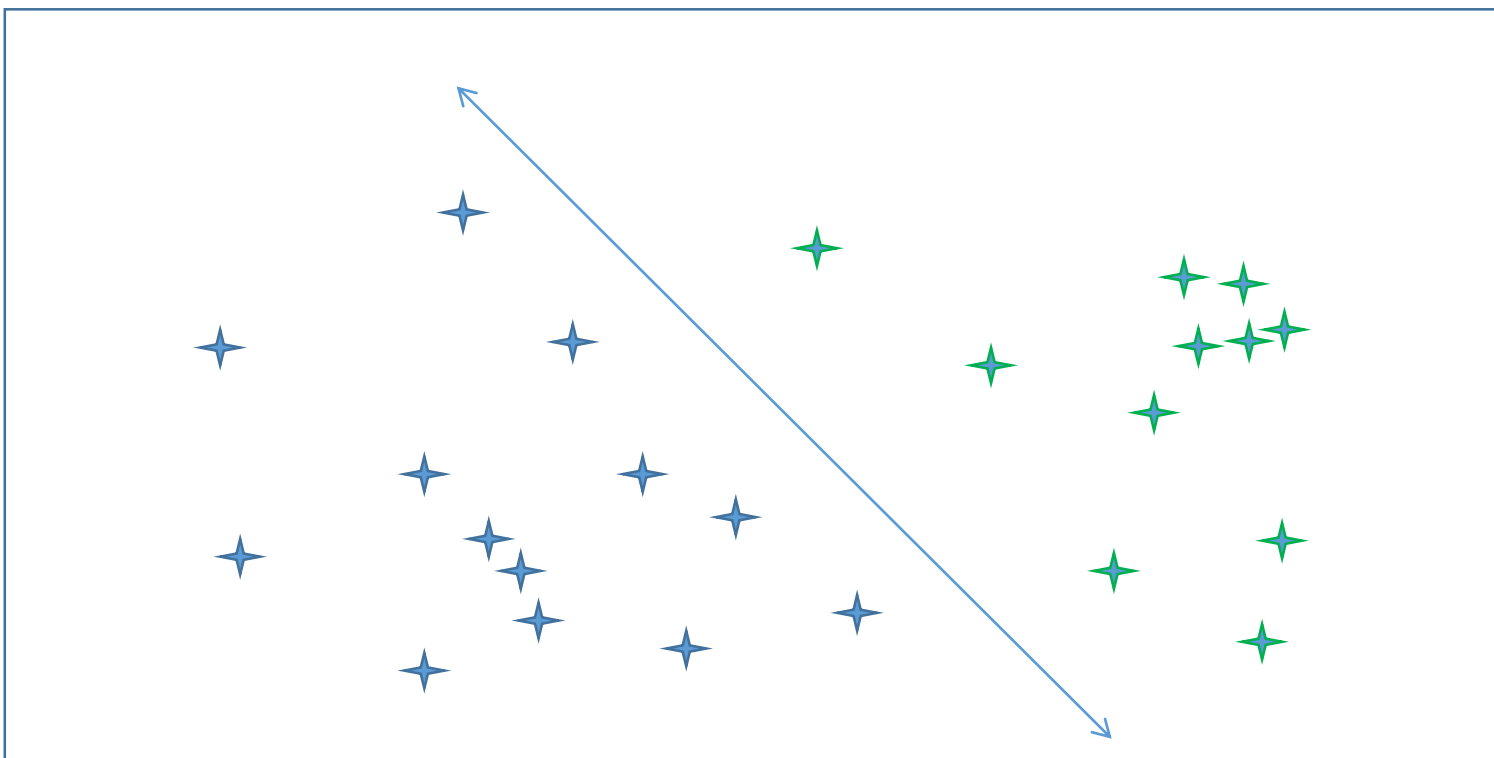
- It's hard to visualize clusters with high dimensionality





# Classification

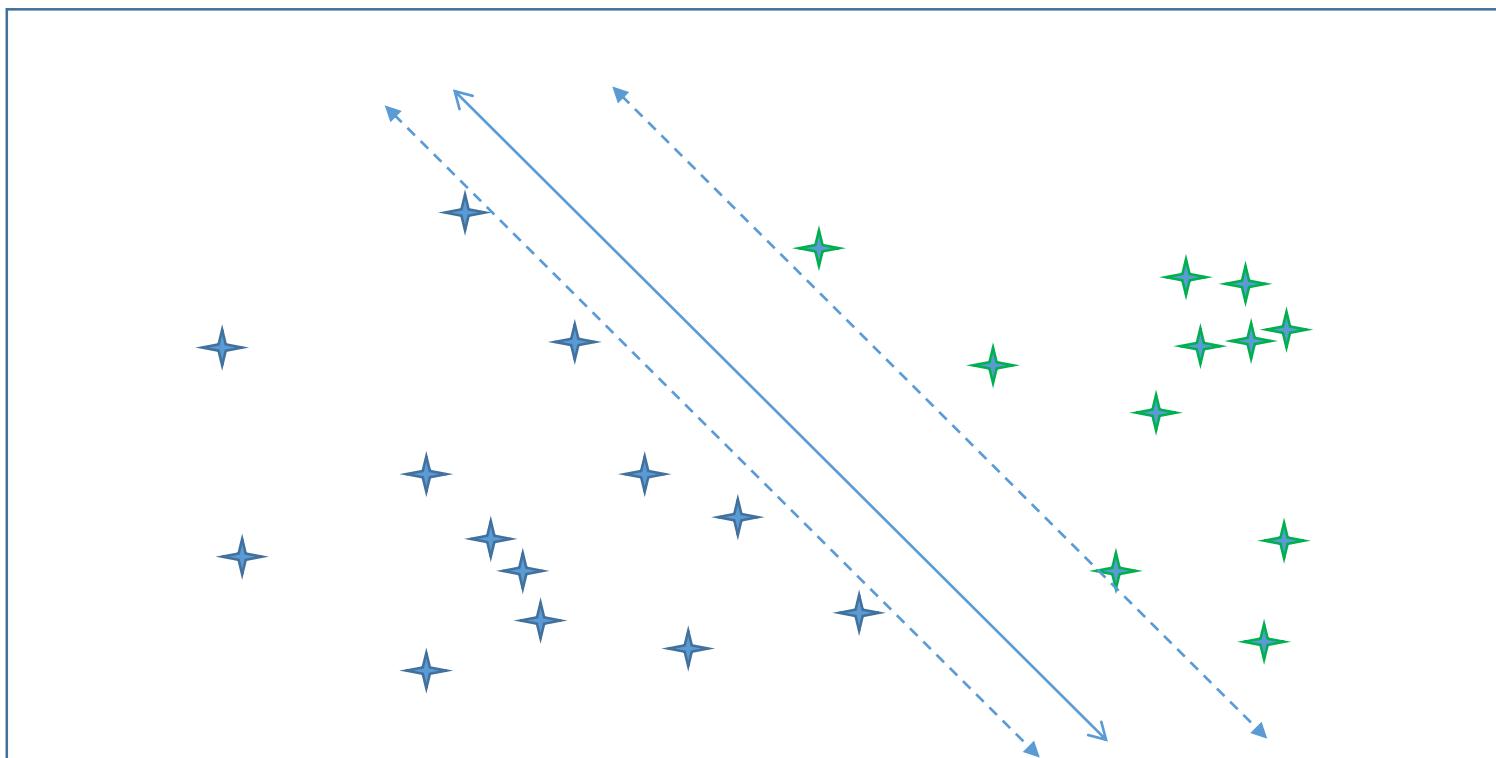
- Regression works well for linearly separable groups





# Classification

- Support Vector Machine works through identifying separable space

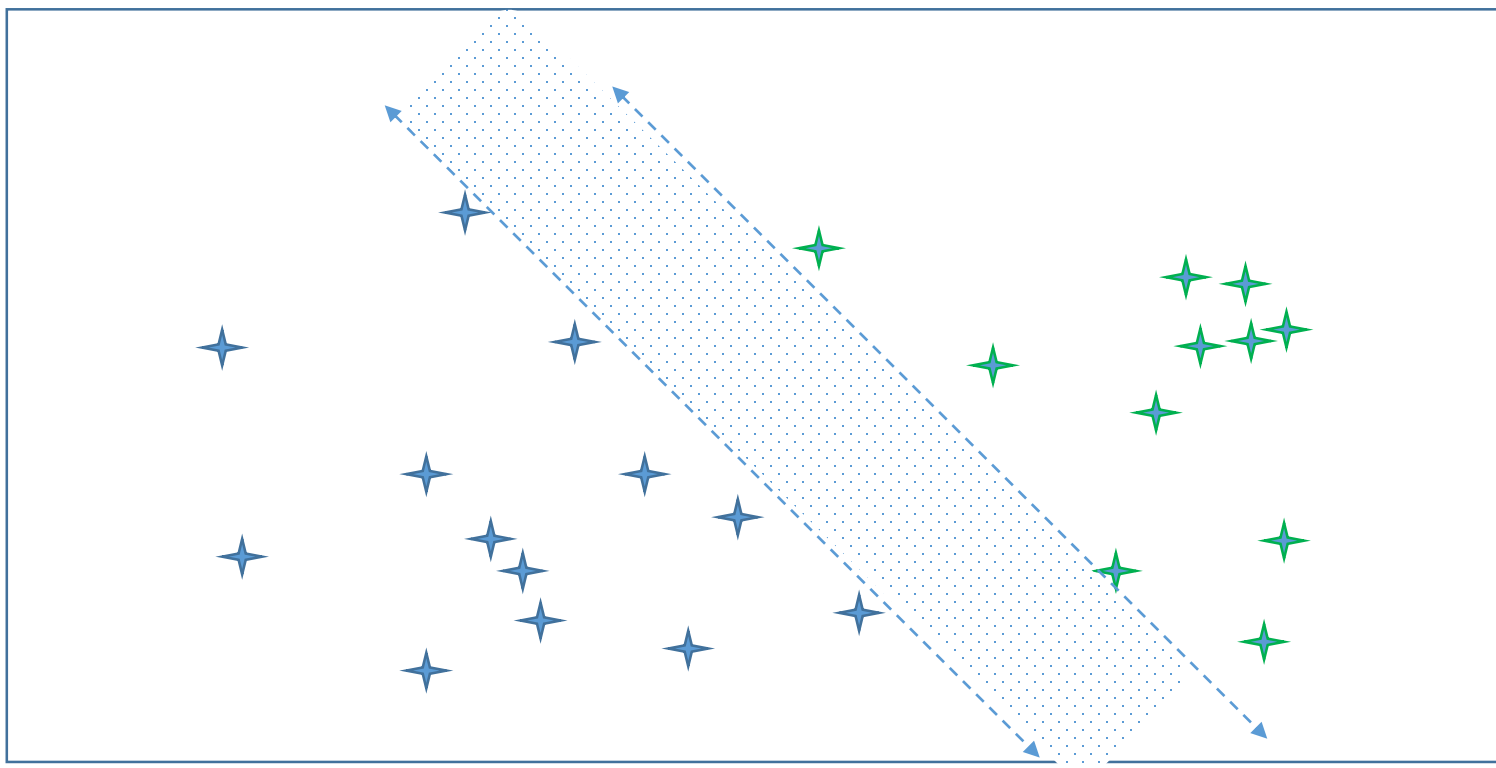






# Classification

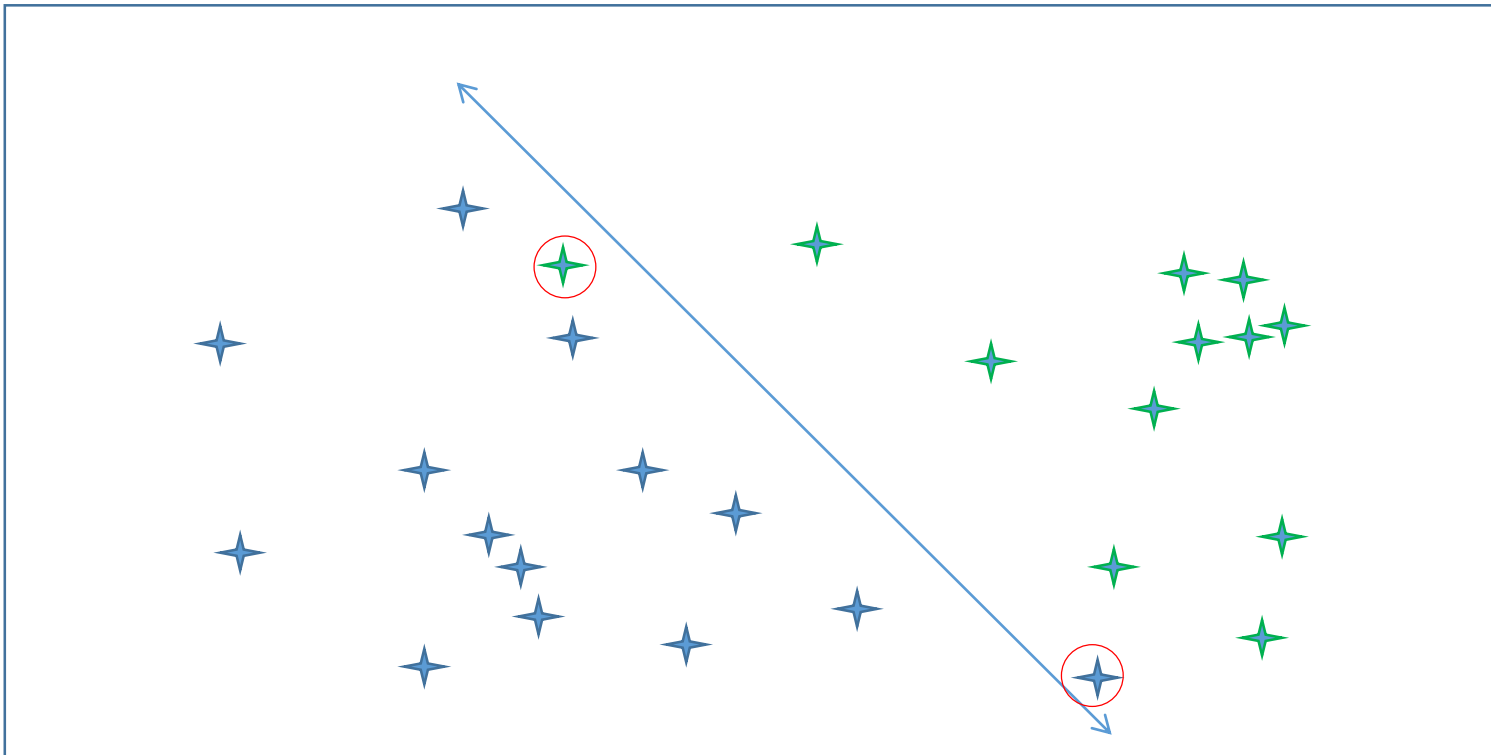
- Support Vector Machine works through identifying separable space





# Classification

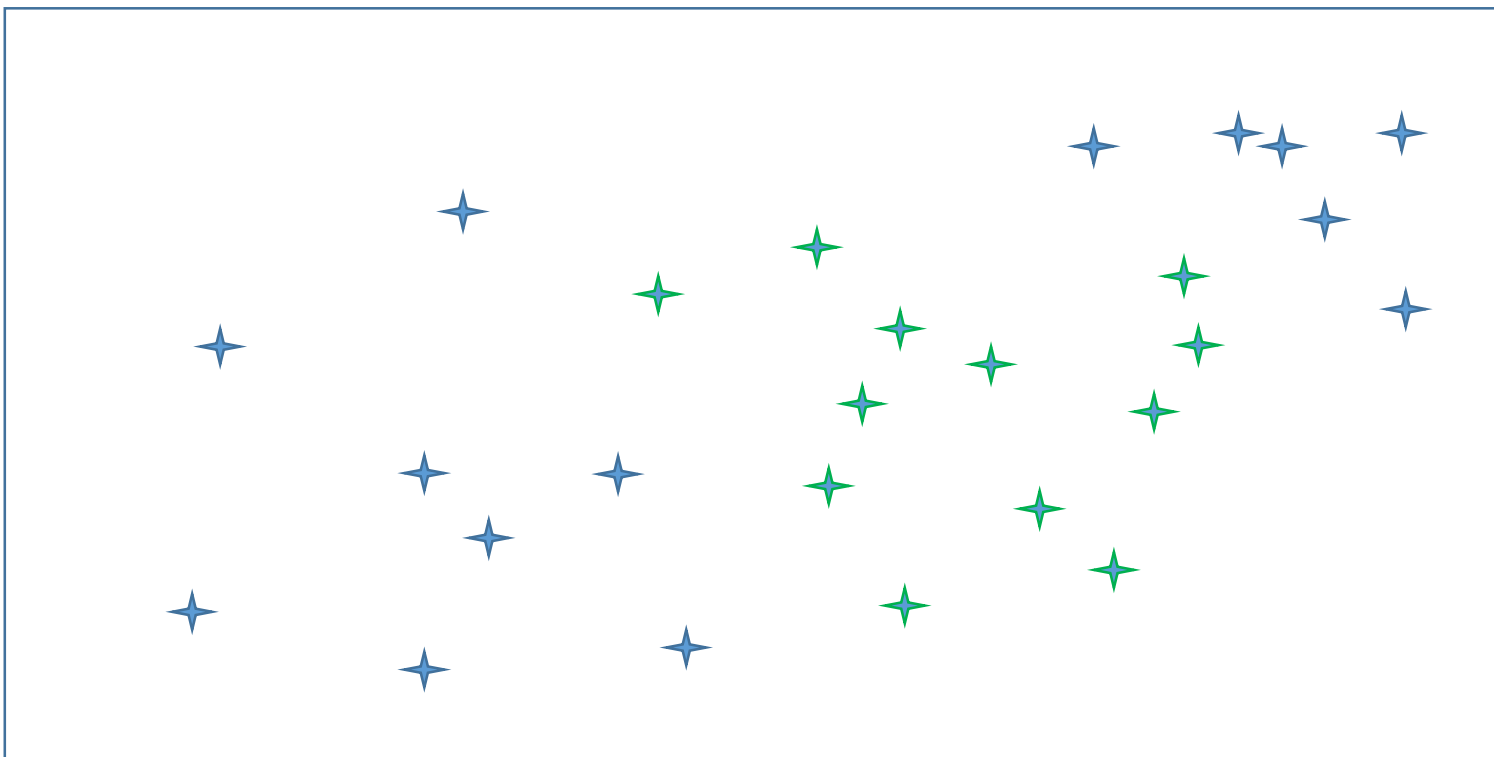
- Newly added members near the boundary may lead to errors





# Classification

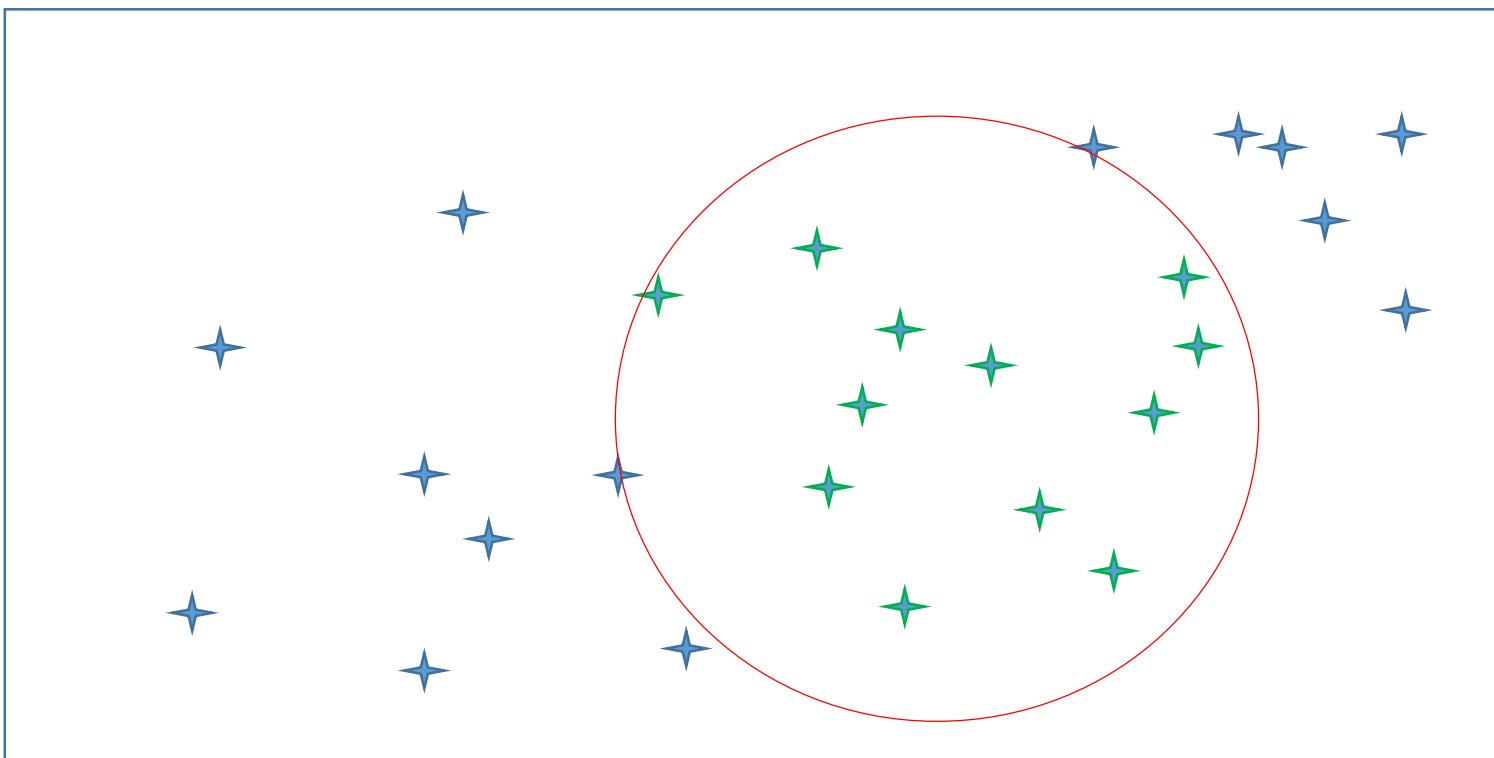
- Other patterns do not lend themselves to linear separation





# Classification

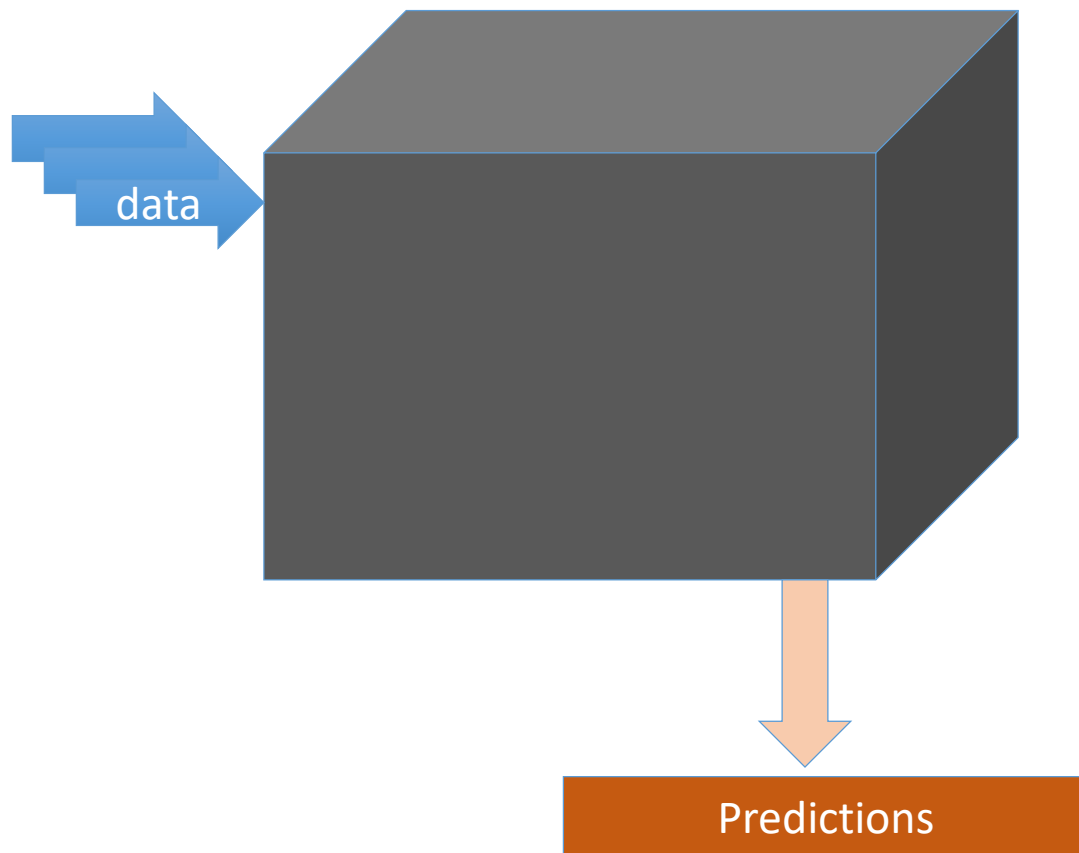
- Kernel functions in SVMs enable non-linear separations





# Neural Nets are Black Boxes

- Can be very accurate compared with other methods
- Very difficult to explain outputs







# Overview of Oracle ML & Advanced Analytics

- Oracle Applications (HCM, EPM, Retail, etc.)
  - Support or enable business functions or operations (**Software as a Service**)
  - Feature “Adaptive Intelligence” or software that has a built-in feedback for learning
  - Can include prebuilt reports and screens
- Oracle Analytics (OAC, DVCS, OBIEE, Essbase)
  - Provide a capability to develop analytics (**Platform as a Service**)
  - Feature machine learning and advanced analytics capabilities for citizen data scientists and professional business analysts
  - Provides the toolsets and frameworks
- Other Oracle
  - Advanced Analytics option to Oracle DB -- Oracle Data Mining, Oracle R Enterprise
  - Autonomous Data Warehouse Cloud Service - Oracle Machine Learning
  - Oracle R Advanced Analytics for Hadoop (ORAAH)
  - Oracle Stream Analytics, Property Graph algorithms, and much more



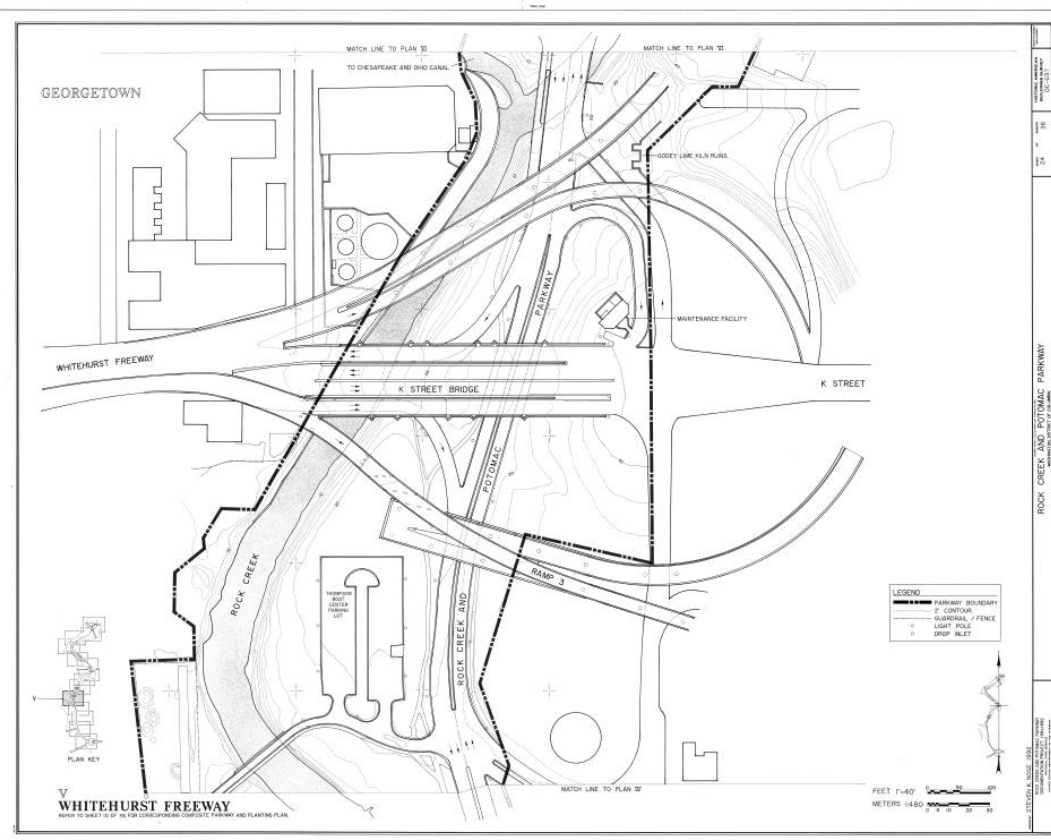
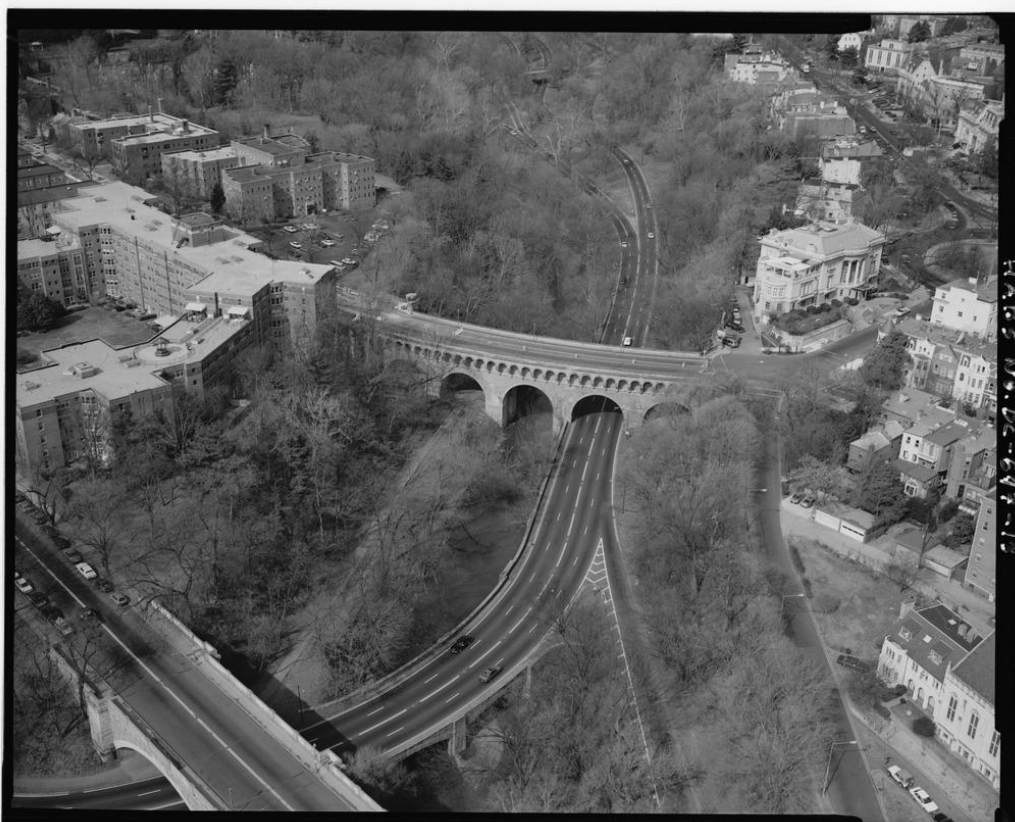
# Major Use Cases and Algorithms

- Predict Lifetime Value of Customers
  - Use regression to project current gross profit contributions into the future
  - Use clustering to group products and customers
  - Use classification to predict likelihood of defection/churn
  - Use decision trees to assign marketing program incentives
- Optimize production processes
  - Use classification to set acceptable run standards
  - Use regression to predict costs of bad quality
  - Use association rules to determine optimal warehouse layout



# Easy and Hard

- Adaptive intelligence is accessible, but requires planning and knowhow







# Clean Data is Essential



Copyright © 2018, vlamiS Software Solutions, Inc.





# Dirty Data is a Pollutant







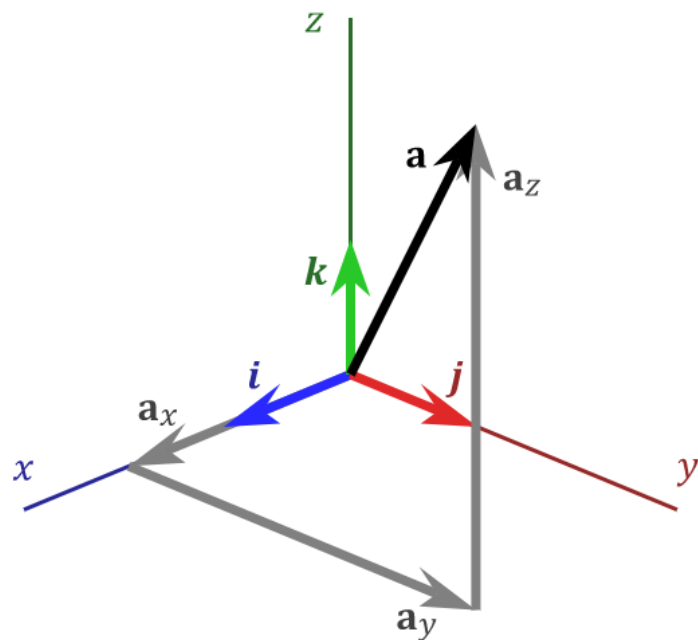
# Bounded vs. Unbounded Domains

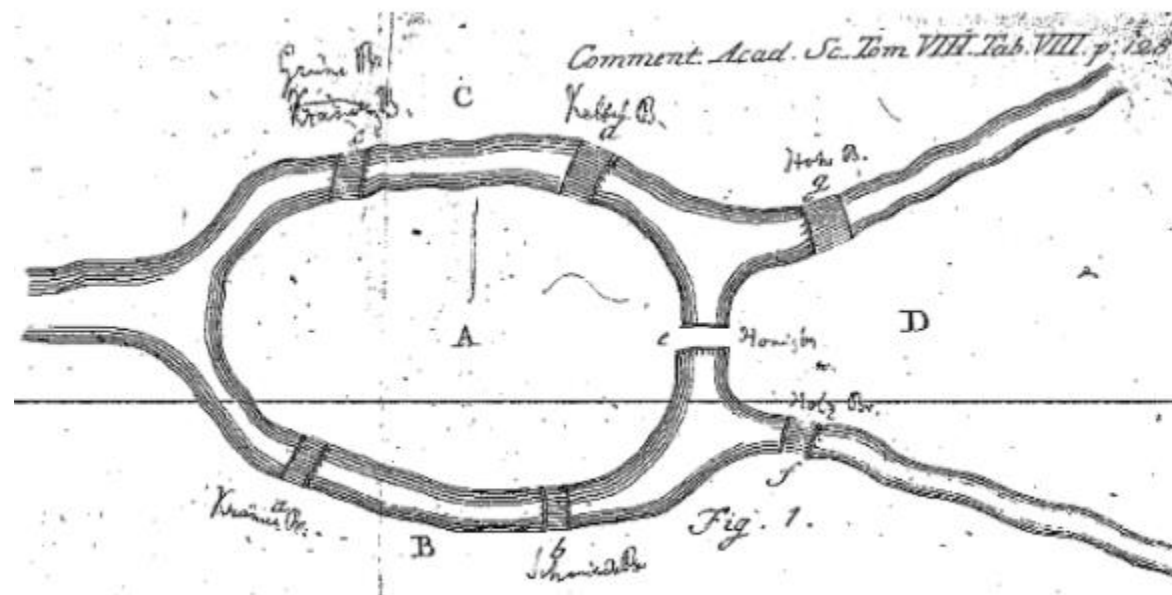
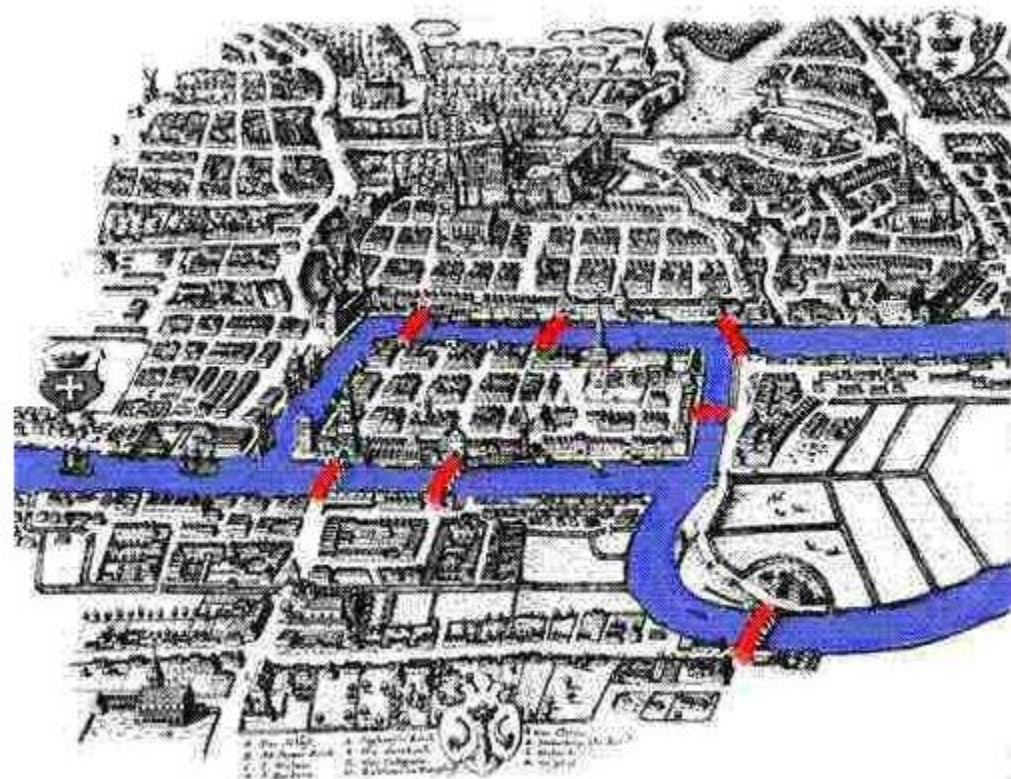
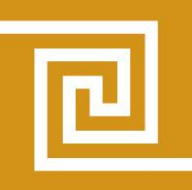
- Bounded “games” like poker, baseball, elections, website A/B testing, etc.
  - Defined rules, time, and results
  - Can use “classic” statistics for prediction
  - Scale space is predetermined
  - Neural nets are excellent for classification exercises in bounded domains
- Unbounded games like the stock market, economic growth, forests, profitability, etc.
  - Significant challenges exist using “classic” statistics
  - Assumptions are both necessary and more important than anything else
  - Scale space is undetermined
  - AI does not do well with unbounded predictions



# Linear Algebra is the new Calculus

- If calculus is the study of motion, change, and forces, linear algebra is the study of the relationships between members of a set as defined by equations.
- Linear Algebra deals with vectors, matrices, transforms, and graphs.





Euler's Figure 1 from 'Solutio problematis ad geometriam situs pertinentis,' Eneström 53 [source: [MAA Euler Archive](#)]



# Frameworks for Thinking about AI & ML

- Stakeholder analysis
- Negotiations/shared interests
- Fiduciary responsibility
- Risk management
- Security
- Data governance and Master Data Management
- Distributive Justice, Ethics, and Moral Philosophy
- Legal framework (HIPAA, FCRA, EU GDPR, etc.)
- Data Mining Frameworks (KDD, CRISP-DM, etc.)
- Complex Adaptive Systems, Systems Dynamics



# European GDPR

## European General Data Protection Regulation

- Pseudonymization and tokenization
- Consent is required
- Data Protection Officer
- Rights to data erasure
- Rights to data portability
- Rights to data access
- Data protection by design and default



# European GDPR

## European General Data Protection Regulation

- Lawful basis for processing
  - the data subject has given consent to the processing of his or her personal data for one or more specific purposes.
  - processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract.
  - processing is necessary for compliance with a legal obligation to which the controller is subject.
  - processing is necessary in order to protect the vital interests of the data subject or of another natural person.
  - processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.
  - processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.



# Bayes Theorem can simply stated

- $P(A | B) = \frac{P(B | A) P(A)}{P(B)}$
- Conditional Probabilities are continuously updated with new results.
- Prior probabilities are updated as posterior probabilities.



# Data Equity and Enterprise Valuation

- The value of data lies in its possession and use.
- The most important data for organizations is their operational data.
- Be *\*very\** careful sharing data with outside organizations.





# Data Value and Data Equity Varies with

- Importance (centrality x connectedness x influence)
- Uniqueness
- Consistency
- Completeness
- Cleanliness
- Freshness
- Refined state
  - Structure
  - Calculated measures
  - Engineered features
  - Aggregations



- Assess the value you are losing through sharing your data
- Assess advantages of growing your analytics capability internally
- Raw data must be processed to provide higher value
- Are all algorithms and engineered features shared with you?
- Are your analytics vendors working with competitors?
- Sharing data means sharing enterprise value



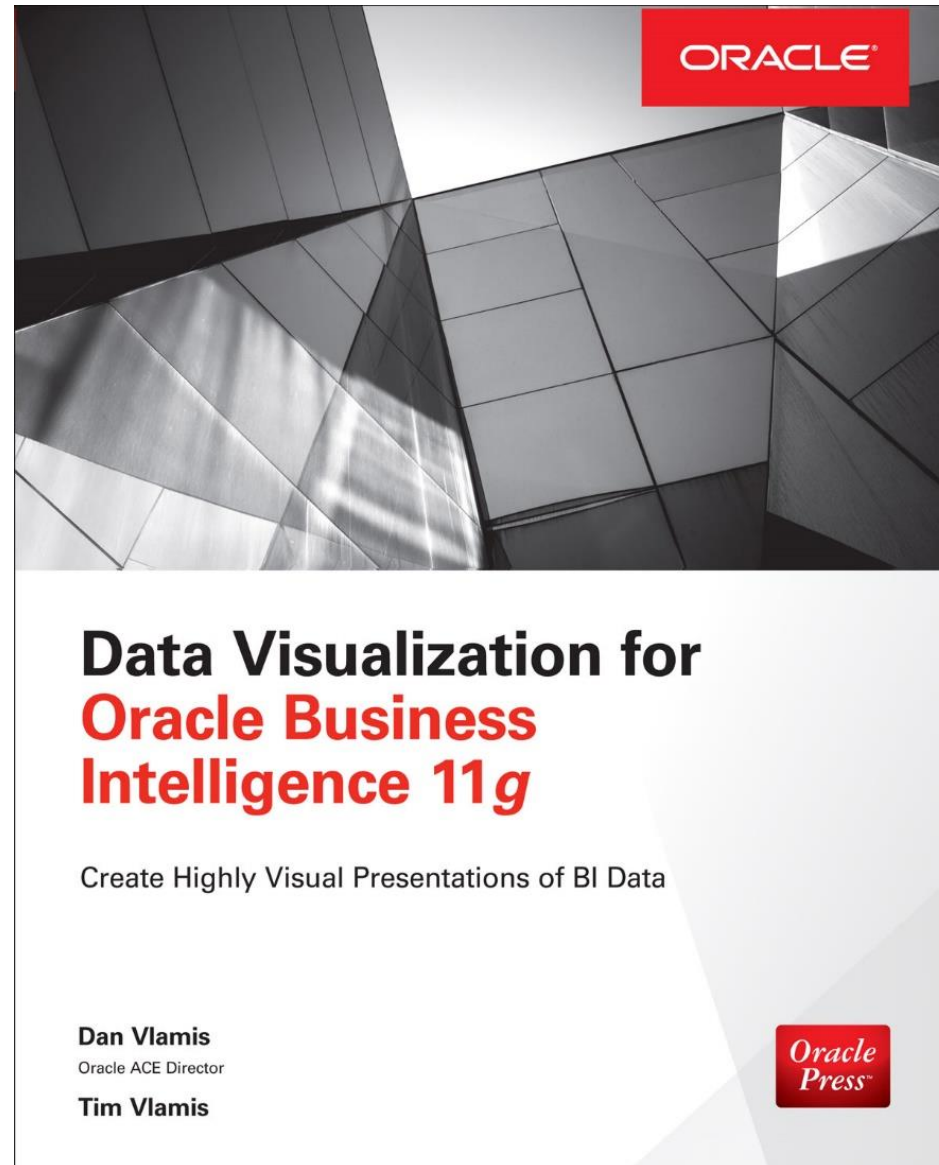
# Summary

- You already encounter and work with Machine Learning every day.
- Learning a little language will help (you already know the concepts)
- Find people who know your business, like data, and can explain math
- Good questions are more important than good answers
- Start



# Drawing for Free Book

Add business card to basket  
or fill out card





# Resources

- [Oracle's Rich Clayton writing about Adaptive Intelligence in CIOReview](#)
- [Is Murder by Machine Learning the New Death by PowerPoint? HBR](#)
- [Trust the Algorithm or Your Gut? HBR](#)
- [Oracle Advanced Analytics on OTN](#) (Oracle Data Mining, Oracle R Enterprise, Oracle Machine Learning)
- [Oracle Data Mining Blog](#)
- [Oracle R Enterprise Blog](#)
- [Oracle Analytics Cloud Machine Learning Blog](#)
- <https://www.kdnuggets.com/>
- <https://cran.r-project.org/>
- <https://www.python.org/>