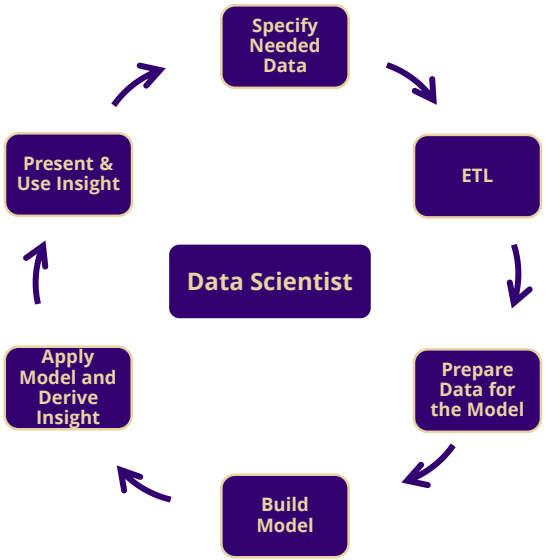
Data Science Cycle

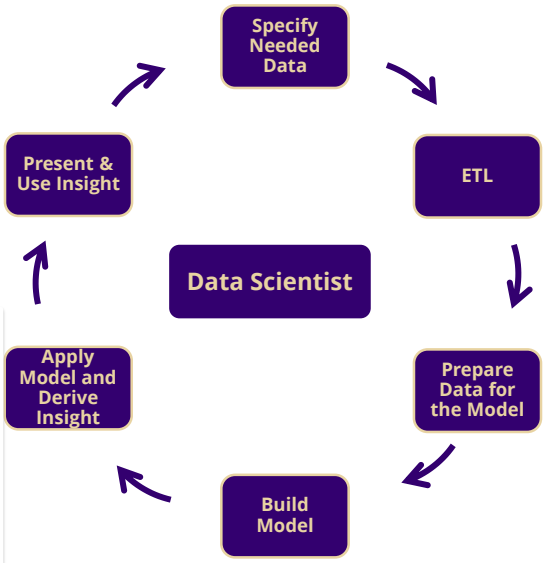
What Data Scientists do

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

Data Science Cycle

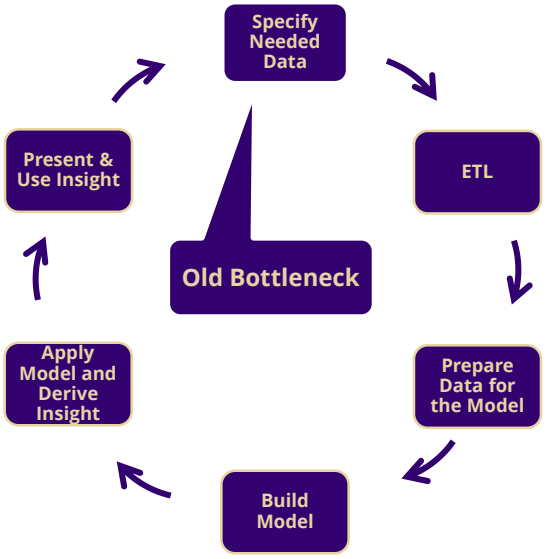


Data Science Cycle

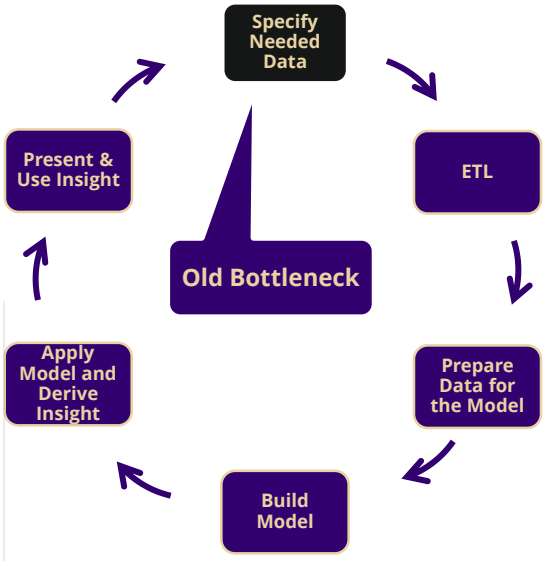


Data Science Cycle
- Data scientist oversees cycle
- Analogies with
  - Cross Industry Standard Process for Data Mining (CRISP-DM)
  - Scientific Method

Data Science Cycle



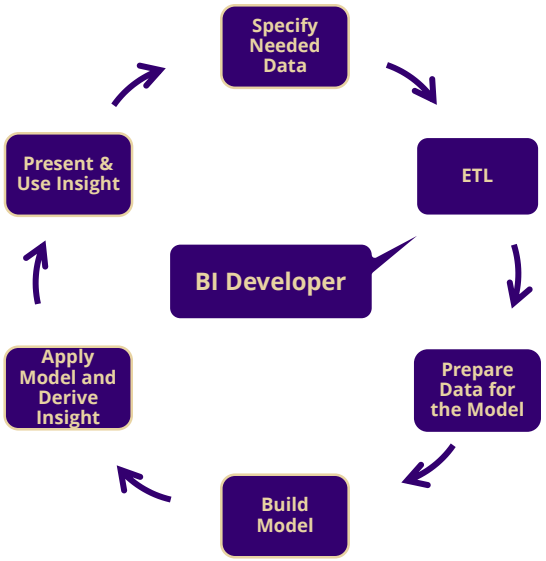
Data Science Cycle



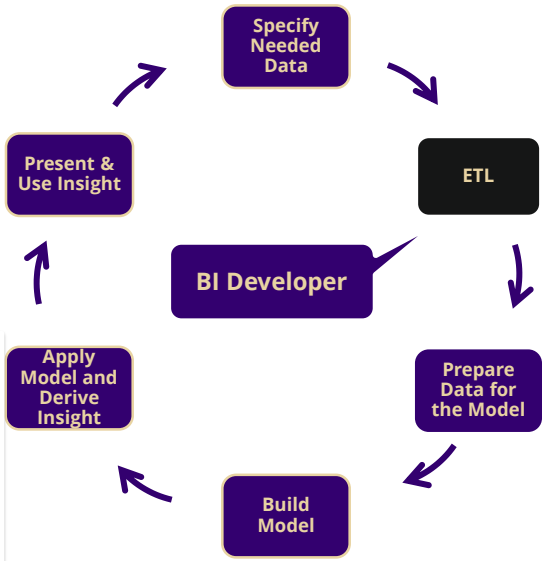
Specify Needed Data

- Data scientist evaluates if data will support a model
- Data scientists rarely create data
- In Design Of Experiment (DOE) data scientists help create data

Data Science Cycle

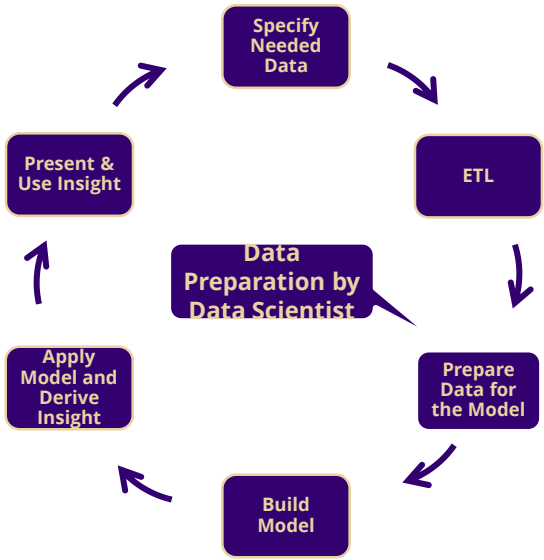


Data Science Cycle

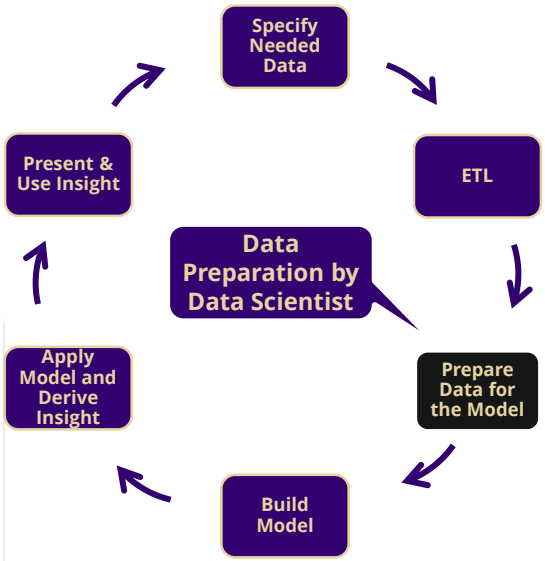


Extract Transform Load (ETL)
- BI developers do ETL
- Transformations in ETL differ from data preparation for models

Data Science Cycle



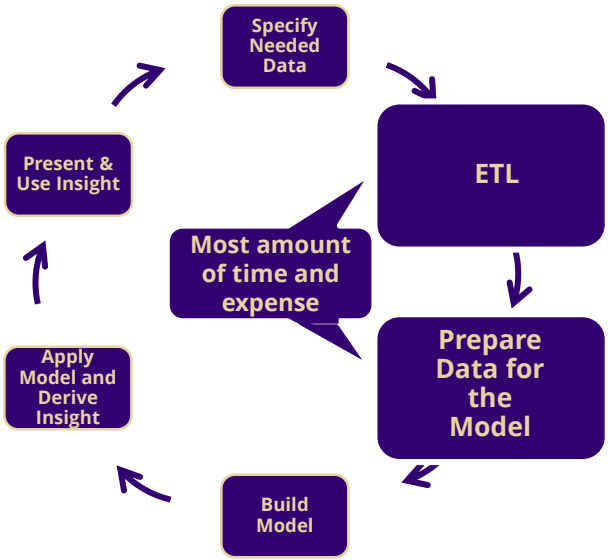
Data Science Cycle



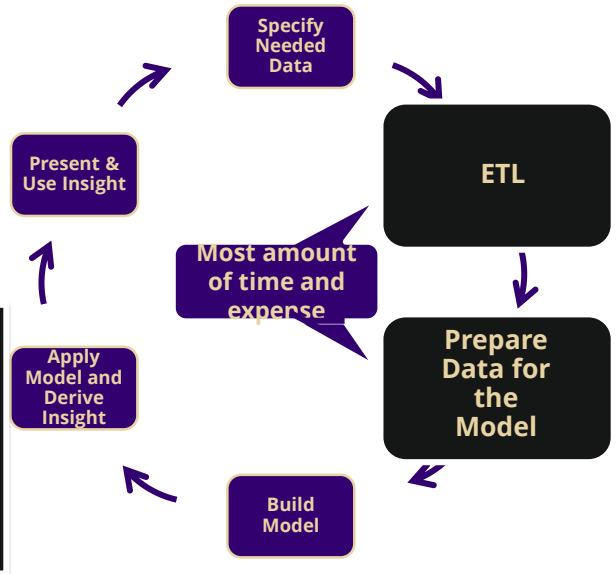
Prepare Data for the Model

- Data Scientists prepare data for modeling
- Data preparation for modeling requires an understanding of the model

Data Science Cycle



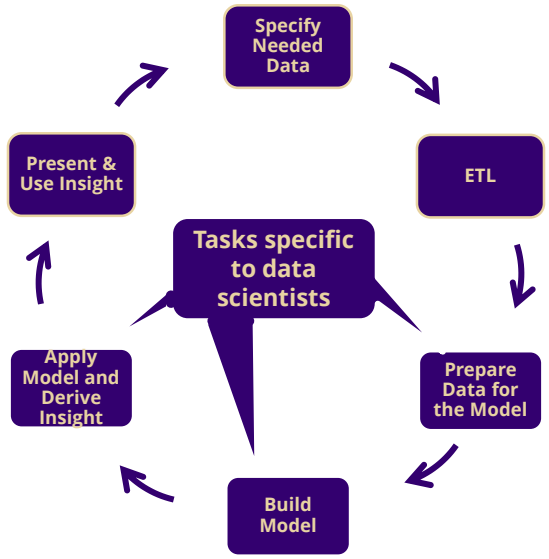
Data Science Cycle



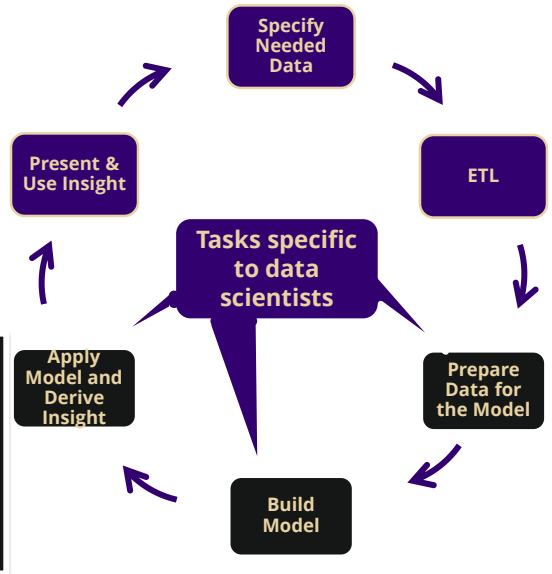
ETL and Data Preparation

- Data access and ETL consume the more time than any other step
- Data preparation consumes most of the data scientists time

Data Science Cycle



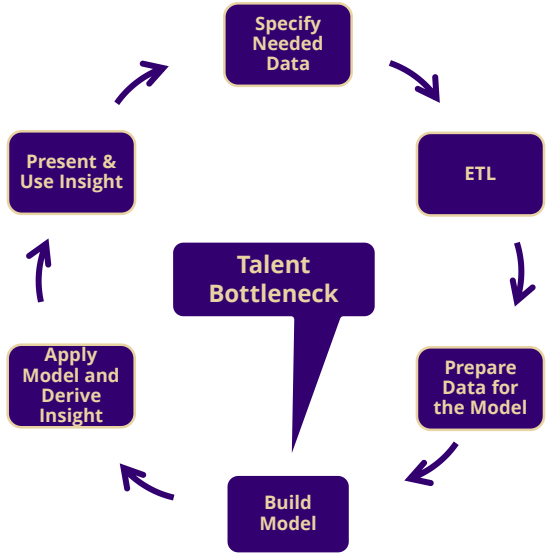
Data Science Cycle



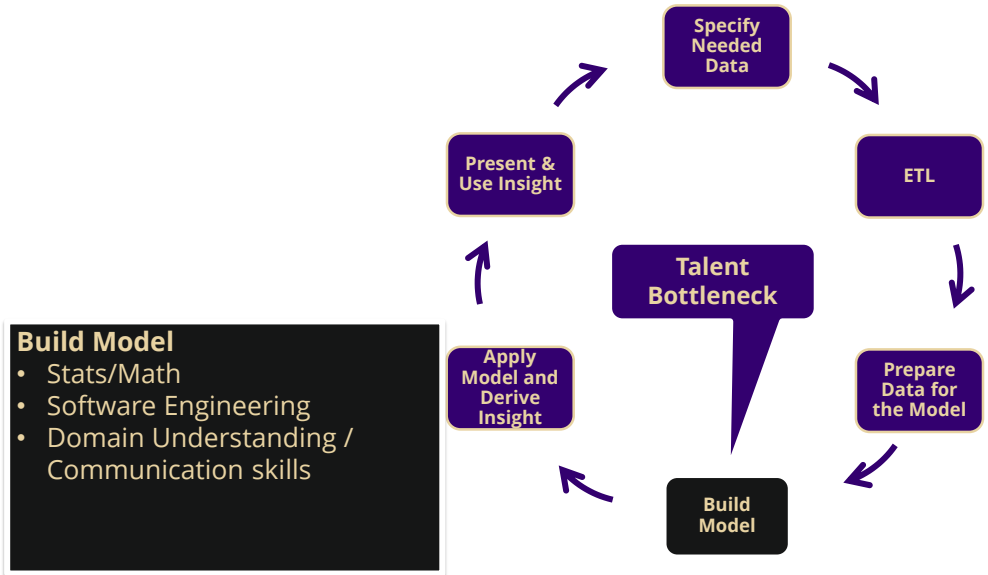
Data Science Tasks

- The data science skill set revolves around creating and applying models

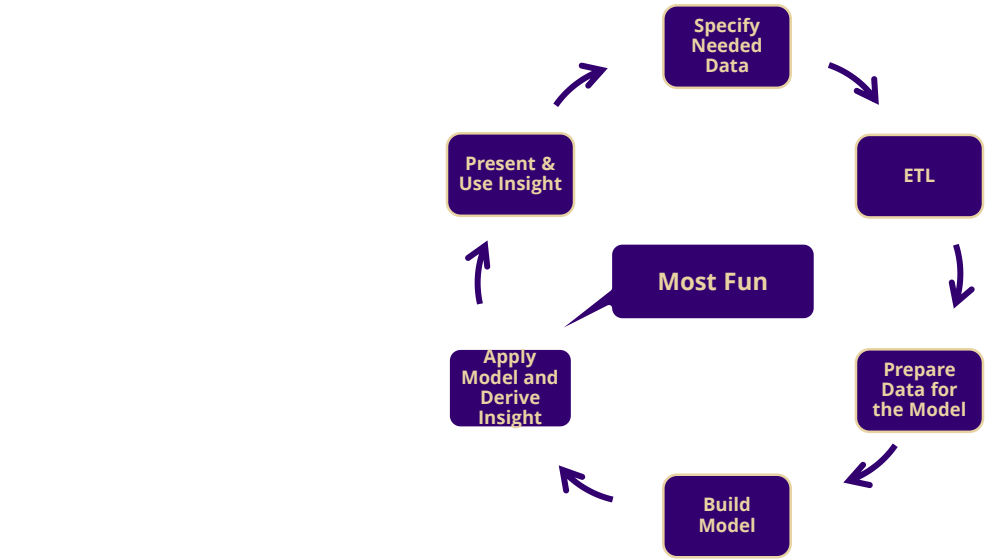
Data Science Cycle



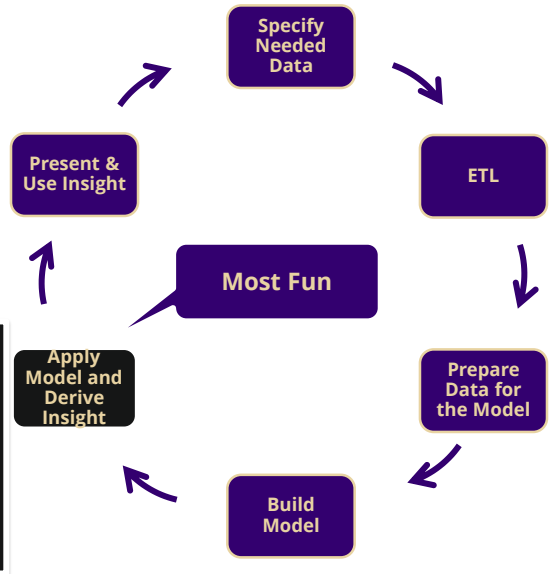
Data Science Cycle



Data Science Cycle



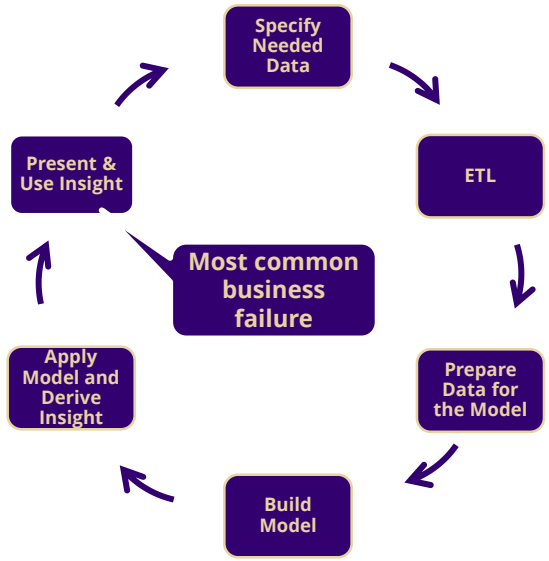
Data Science Cycle



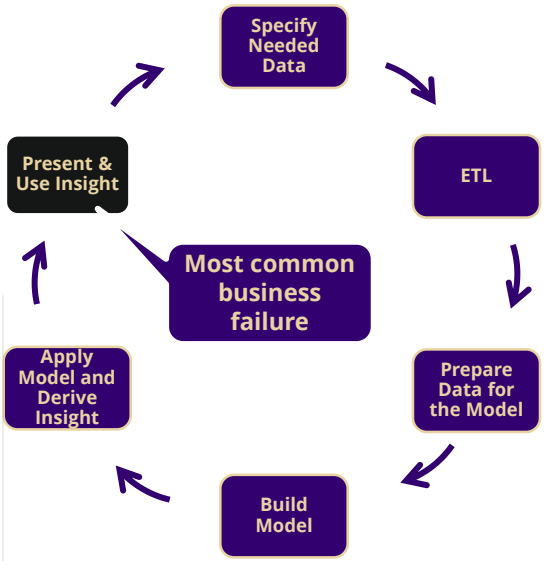
Apply Model and Derive Insight

- Applying the model is the most fun

Data Science Cycle



Data Science Cycle



Present & Use Insight
- Often overlooked
- Often underestimated
- Get result to right person at right time
- Last mile of predictive analytics

Data Science Cycle

What Data Scientists Do

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

The Four Analytics

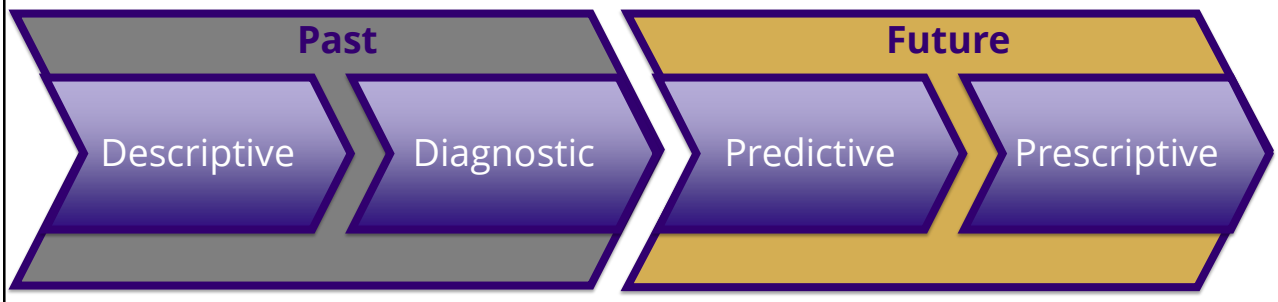
What Data Scientists do

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

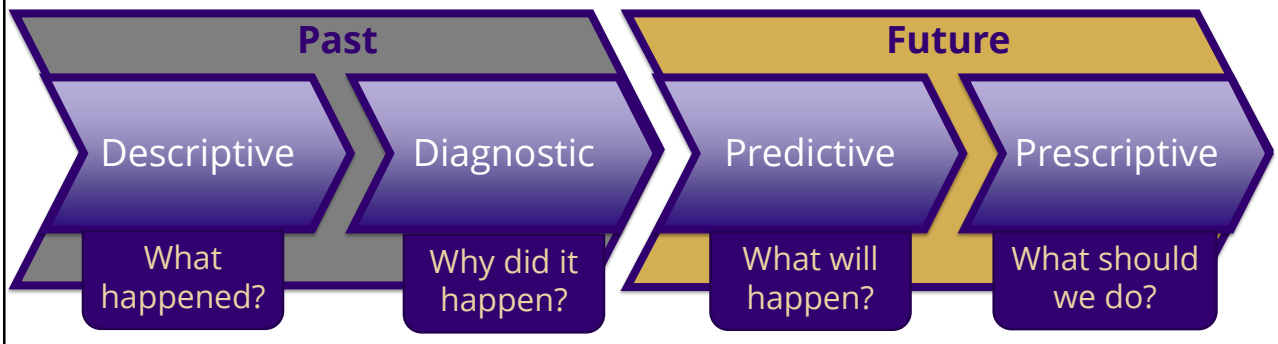
The Four Analytics

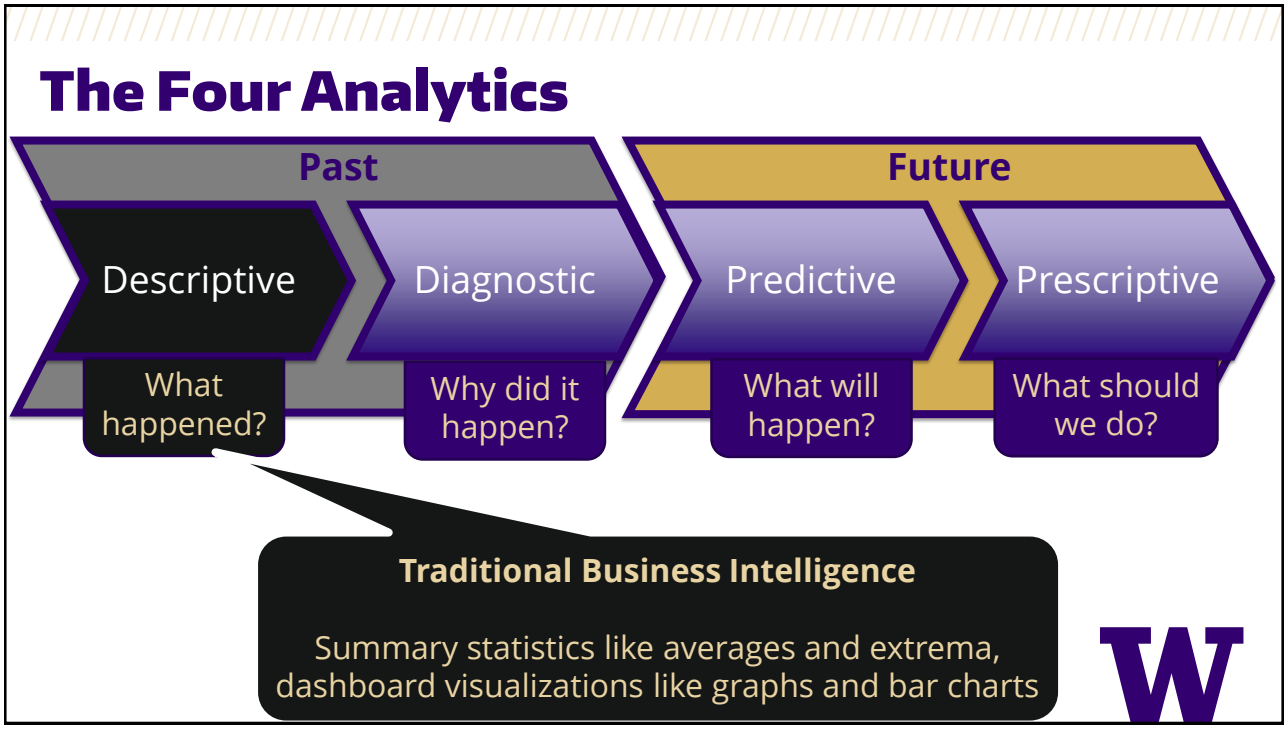
References:
http://insights.principa.co.za/4-types-of-data-analytics-descriptive-diagnostic-predictive-prescriptive
https://dupress.deloitte.com/dup-us-en/topics/analytics/five-types-of-analytics-of-things.html
http://www.rosebt.com/blog/descriptive-diagnostic-predictive-prescriptive-analytics
http://www.gartner.com/binaries/content/assets/events/keywords/catalyst/catus8/2017_planning_guide_for_data_analytics.pdf

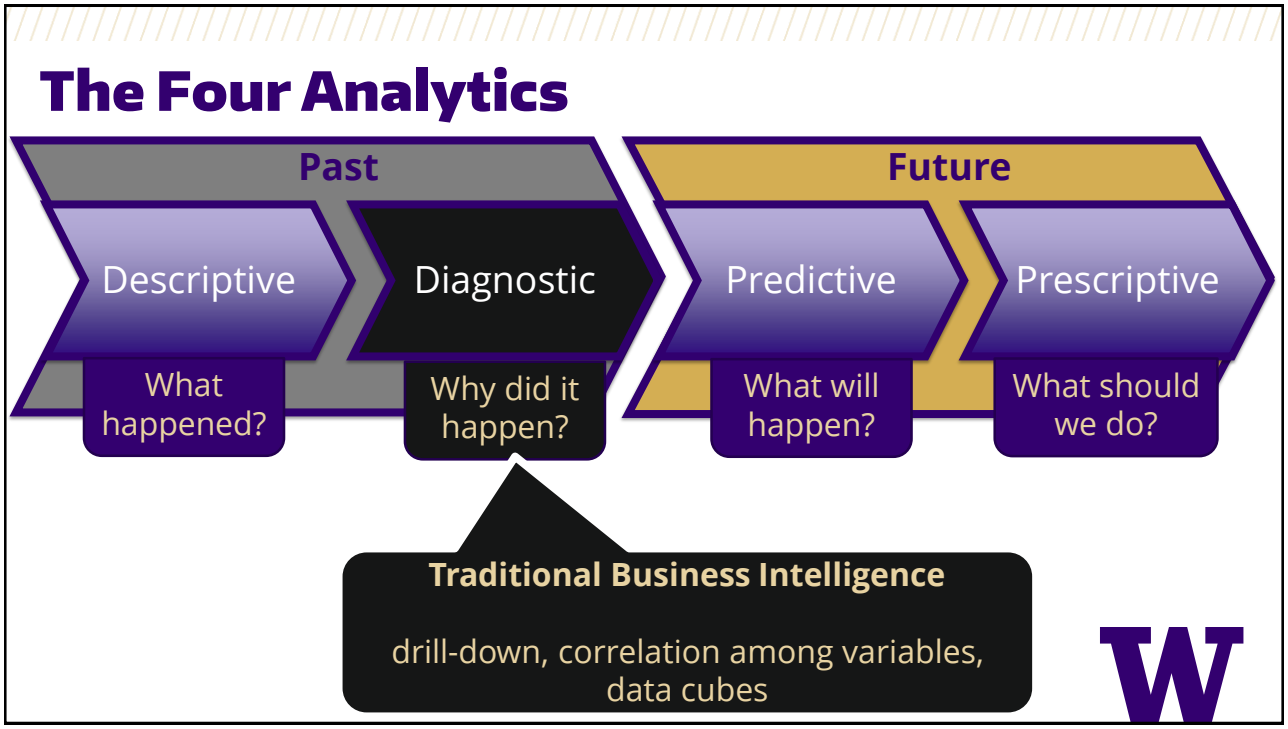
The Four Analytics

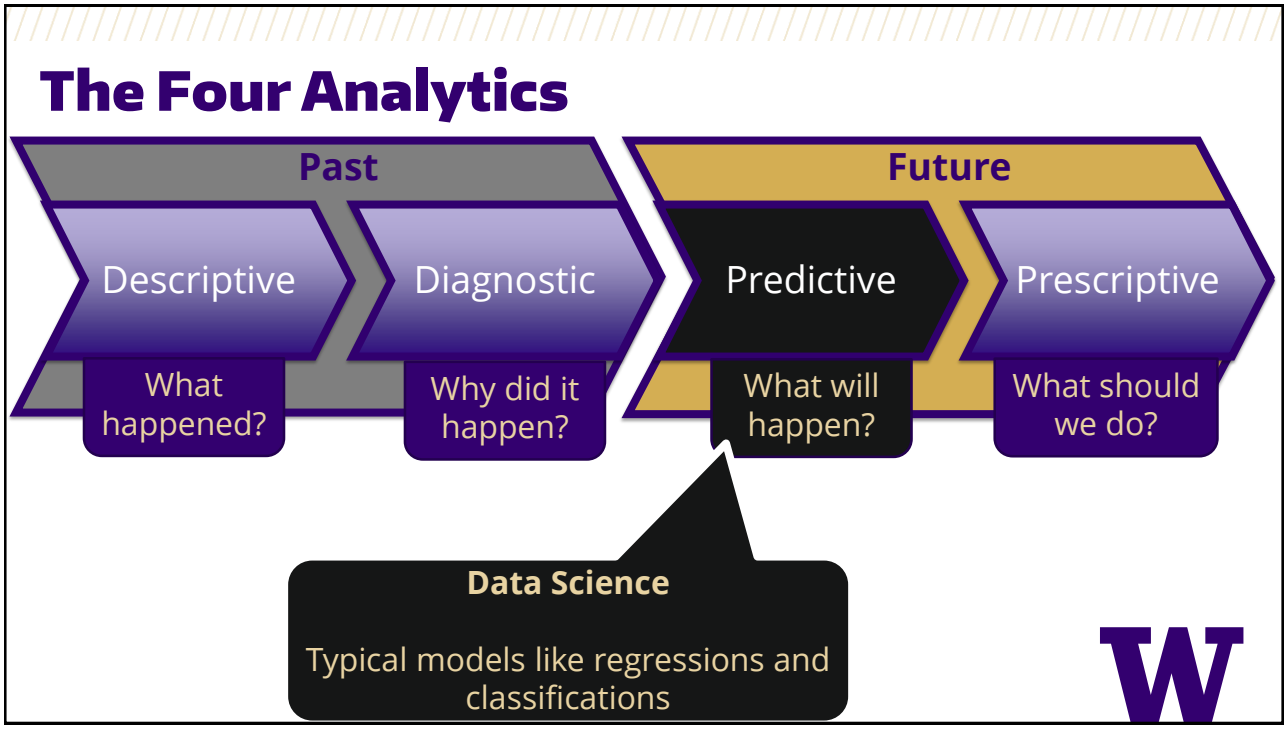


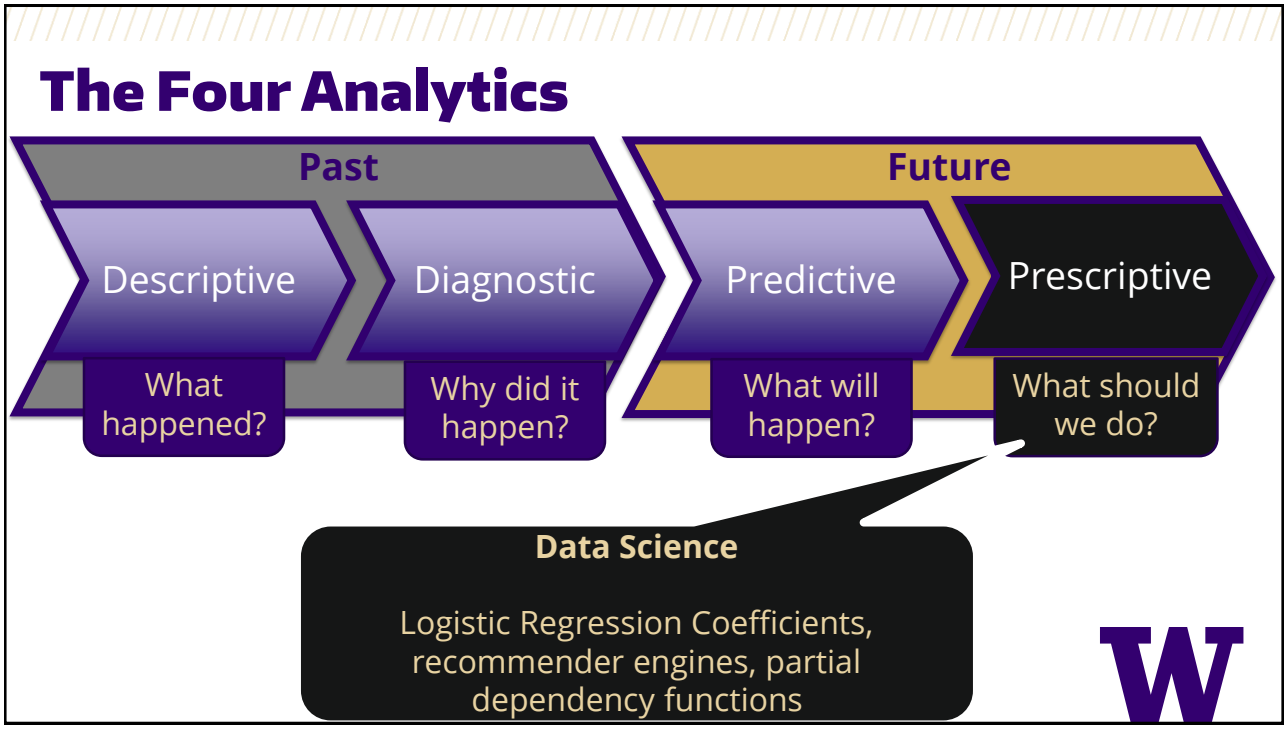
The Four Analytics




The Four Analytics
Past
Future
Descriptive
Diagnostic
Predictive
Prescriptive
What happened?
Why did it happen?
What will happen?
What should we do?
Traditional Business Intelligence
Summary statistics like averages and extrema, dashboard visualizations like graphs and bar charts


The Four Analytics
Past
Future
Descriptive
Diagnostic
Predictive
Prescriptive
What happened?
Why did it happen?
What will happen?
What should we do?
Traditional Business Intelligence
drill-down, correlation among variables, data cubes


The Four Analytics
Past
Future
Descriptive
Diagnostic
Predictive
Prescriptive
What happened?
Why did it happen?
What will happen?
What should we do?
Data Science
Typical models like regressions and classifications


The Four Analytics
Past
Future
Descriptive
Diagnostic
Predictive
Prescriptive
What happened?
Why did it happen?
What will happen?
What should we do?
Data Science
Logistic Regression Coefficients, recommender engines, partial dependency functions

The Four Analytics

What Data Scientists Do

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON