



# Leveraging Educational Data Mining Techniques to Determine Factors Impacting Military- Connected Students

Andrew Berning Ph.D. The Renaissance Institute

Ryan Baker Ph.D. University of Pennsylvania

MCEC National Training Seminar

July 24, 2018



# The Research Question

Can big data and data-intensive modeling support a school district in successfully increasing achievement and efficiency in the lives of PK-12 learners?

# Learning Analytics

- Learning Analytics is the science of the measurement of data about learners and the learning environment/system to visualize and optimize the complex conditions in a typical school system.
- Predictive Analytics is a type of learning analytics where we predict future outcomes from data today and determine why an outcome is likely, so that we can influence that outcome.
- Predictive Analytics now used with increasing frequency to study K-12 student success (Bowers et al., 2012; Knowles, 2015; Baker et al., under review).



# K-12 Analytics Research Consortium

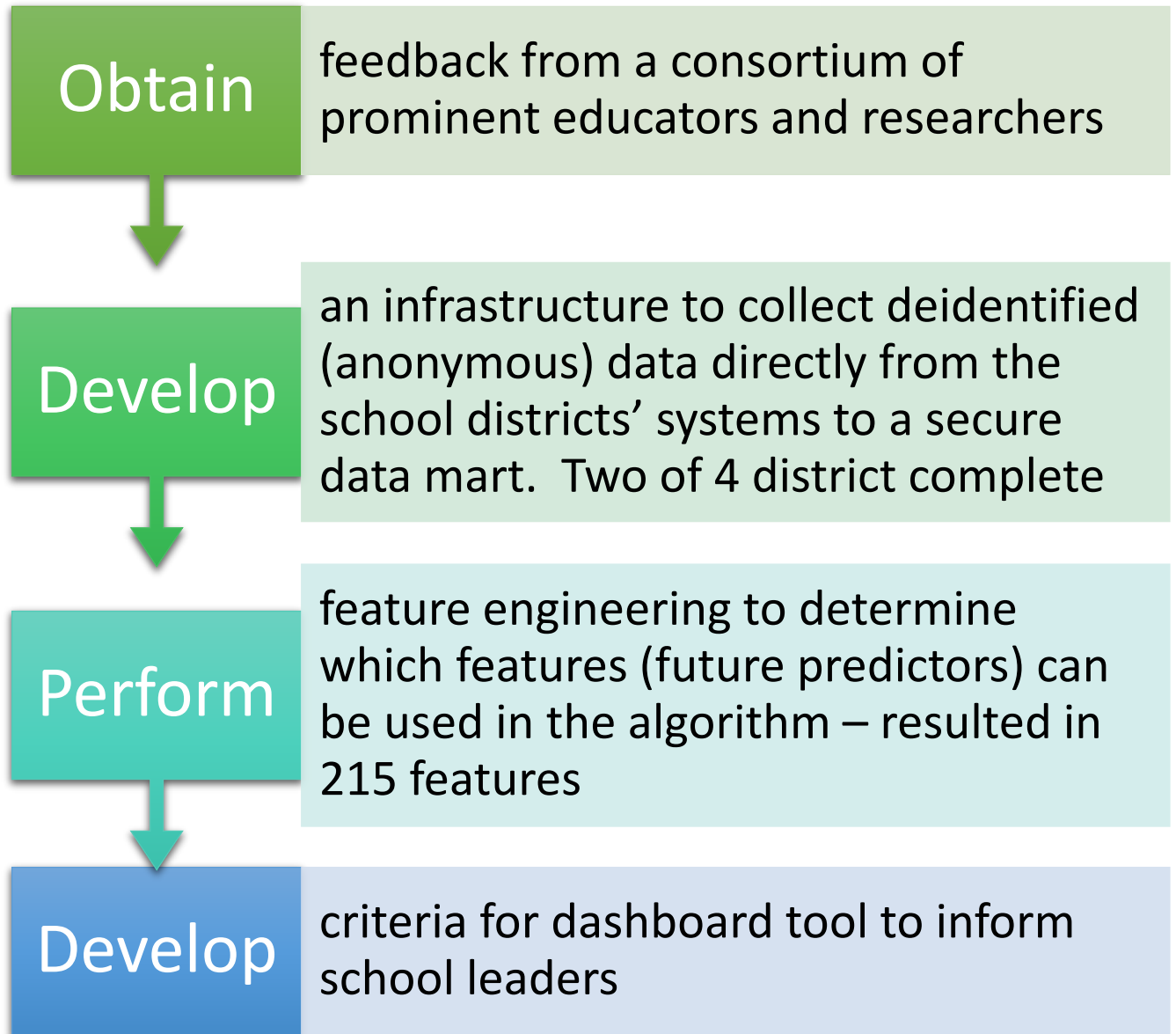
- K-12 schools in Texas working with the Renaissance Institute
- Insight into issues of critical importance to K-12 leaders
  - Feedback on the impact of technology
  - Feedback on district initiatives
  - Predicting school success
  - Predicting student success
  - Predicting school failures

# The MCEC Study

Funded by  
The AT&T  
Foundation

- A discovery project to identify if modern data mining techniques can be leveraged to support military-connected students
- Included four (4) Texas K-12 school districts totaling over 60,000 students
- Predictive analytics methods used to predict future success for all students
  - Will a student graduate from high school?
  - What will a student's SAT score be?
- Can these techniques be used to identify features unique to military-connected students?
  - How do military-connected students differ from the broader population of PK-12 students?

# The Process



# 215 Features Engineered from School District's Student Information System

- Included combinations of
  - Courses taken
  - Grades
  - Attendance
  - Discipline
- Did not include demographic data
  - Race
  - Gender
  - Income
  - Address
  - Feeder pattern

# Methods

- Designed and developed a secure data mart and collection strategies
- Developed data dictionary
- Gathered three years of deidentified data directly from district's Student Information System
- Cleaned data
- Analyzed three years of data
- Derived models to predict graduation and SAT



# Results



# Ability to predict graduation (AUC ROC) among military-connected students

Model developed with data from	District 1	District 2
<b>Military-connected</b>	<b>0.75</b>	<b>0.76</b>
<b>Non-military-connected</b>	0.69	0.74
<b>Chance</b>	0.50	0.50

AUC ROC indicates model's ability to differentiate a student who will graduate from a student who will not graduate

# Ability to predict SAT (Spearman Rho) among military-connected students

Model developed with data from	District 1	District 2
<b>Military-connected</b>	<b>0.53</b>	<b>0.51</b>
<b>Non-military-connected</b>	0.44	0.51
<b>Chance</b>	0.00	0.00

Rho indicates correlation between model prediction and actual score – Rho controls for non-normality in data

# Model Findings

Models include findings such as

- Average semester grade is a strong positive predictor of SAT and graduation (more in District 1 than District 2, but true for both)
- Military-connected students perform better on SAT if they have been in district longer
- Average final grade in mathematics and science are more predictive for military-connected students than non-military-connected students

# Model Findings

- Discipline and attendance features were predictive of success, although codes vary among districts
- It is possible to predict eventual graduation and SAT scores from Junior year data for both military and non military-connected students
- The models can be used to focus interventions and to support all students

# Next Steps

Confirm

model with additional school districts. Two additional underway at this time

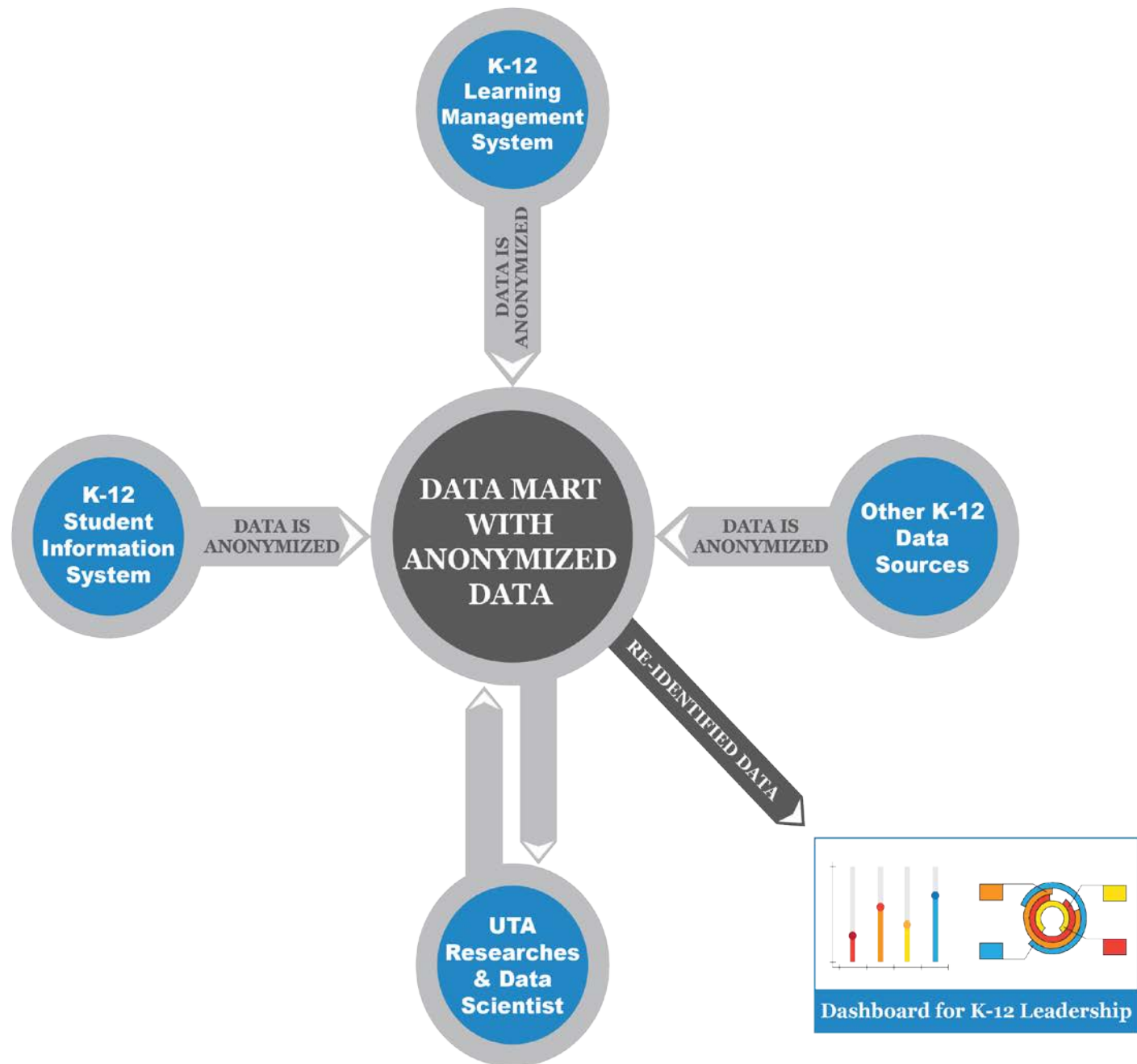
Add

Additional features including learning management system data and post-secondary success data

Develop

Develop feedback tool/dashboard to re-identify data and provide insight to school leadership

# Proposed Tool Framework



# Conclusion

- The Military Student Identifier is a critical data element to support all students
- Modern data mining techniques can be leveraged to identify predictors of success
- These predictors are actionable for interventions
- Real-time, predictive tools have been developed to inform school leaders



# Disclaimer

- Models presented here give no evidence with regard to causality
- Further work, with additional school districts, is needed to verify models

# Special Thanks to The AT&T Foundation

- AT&T is committed to advancing education, strengthening communities, and improving lives.
- Supporting education has been a major focus for AT&T for more than a century and education is our key philanthropic focus.
- We are driven to help students succeed -- in school, in the workforce, and in life.

# Questions and Comments

Andy Berning  
[andy@teacherinfo.com](mailto:andy@teacherinfo.com)



Ryan Baker  
[ryanshaunbaker@gmail.com](mailto:ryanshaunbaker@gmail.com)