R for Equity-Focused Improvement:

Using R Markdown, openxlsx, and ggplot to Empower Education Leaders in California

Authors: Adriana Rodriguez (arodriguez@edanalytics.org) & Robert Little (rlittle@edanalytics.org)

Education Analytics INC.

CORE DISTRICTS

Building a Data System for Schools

Benefits of a Data System:
- Automated data cleaning...
  so we can spend more time on analysis!
- Functionalized business rules
- Standardized data, easy for researchers to use

Apache Airflow

We use Python & Apache Airflow for Automation:
- Able to integrate R and SQL
- Clear visualizations of the data pipeline
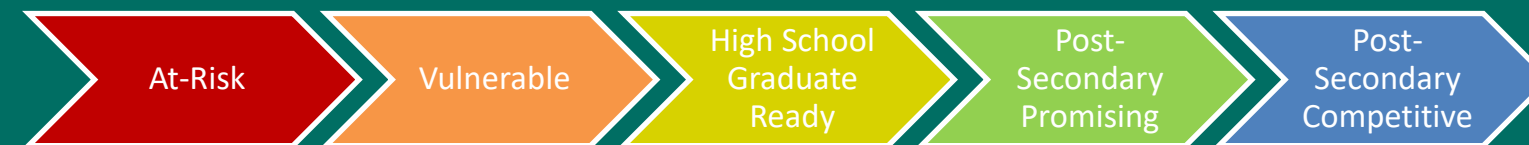
Systematizing Messy Data

Districts have different student information systems and therefore different data:
- Common data issues are forced-fixed
- Automated QC allows for files to fail if it has bad data
- Shiny app created to help identify red flags

The Breakthrough Success Community (BTSC)

Goal: To improve 9th grade on-track rate for high school graduation and college & career success

At-Risk → Vulnerable → High School Graduate Ready → Post-Secondary Promising → Post-Secondary Competitive
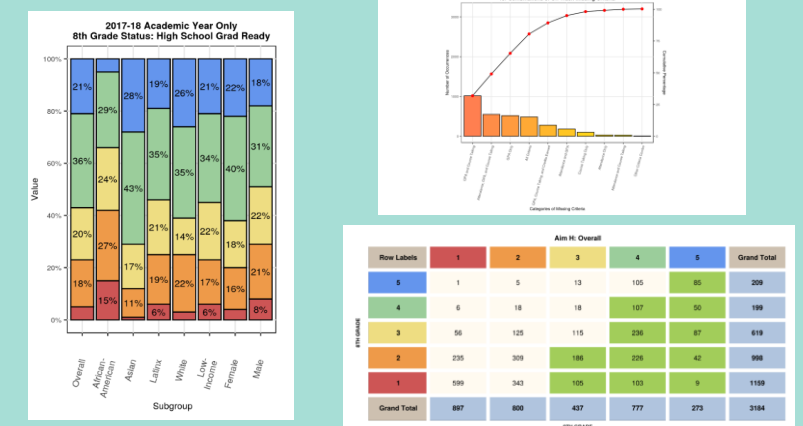
On the ground-floor...

Educators test out best-learning practices and techniques

...On the back-end

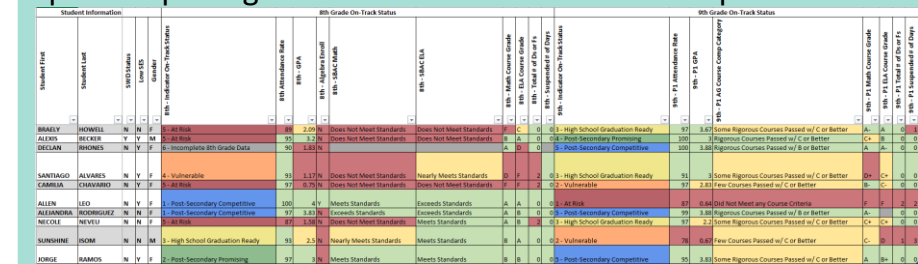EA provides periodic analytics and reports that highlight student progress

School Reports

HTMLs created for district + school level summaries

Student Reports

Openxlsx package creates color-formatted excel spreadsheets

Planning for a Dashboard

Engineering the BTSC process:
- Automating the creation of the BTSC 5-bucket checklist score
- Storing the information in the data system
- BTSC as part of the general CORE database ecosystem