DairyCoPilot - Automated Data Compilation and Analysis Tools for DairyComp Data Assets

Srikanth Aravamuthan, Dörte Döpfer, and Emil Walleser

Department of Medical Sciences, School of Veterinary Medicine, University of Wisconsin, Madison 53706

walleser@wisc.edu

Dairy Farm Management

- Data-driven decisions
- Successful farms record and analyze data
  - Financial
  - Input
  - Production/health
- Animal health
  - Need data and tools

DairyComp

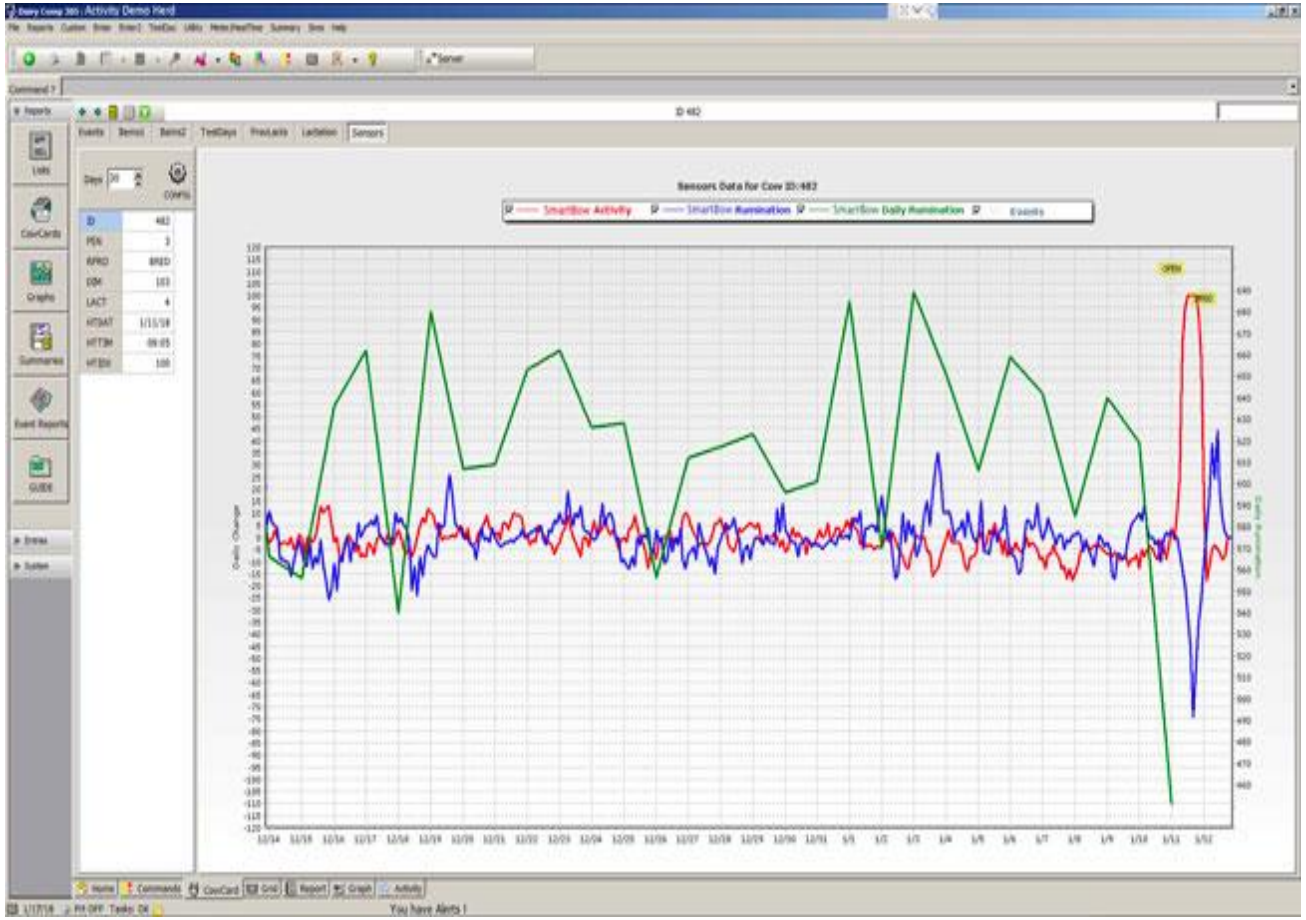
- Dairy management software
- Interconnected apps
- Most popular software

Strengths

- Large user base
- Multiple integrations
- User customizability

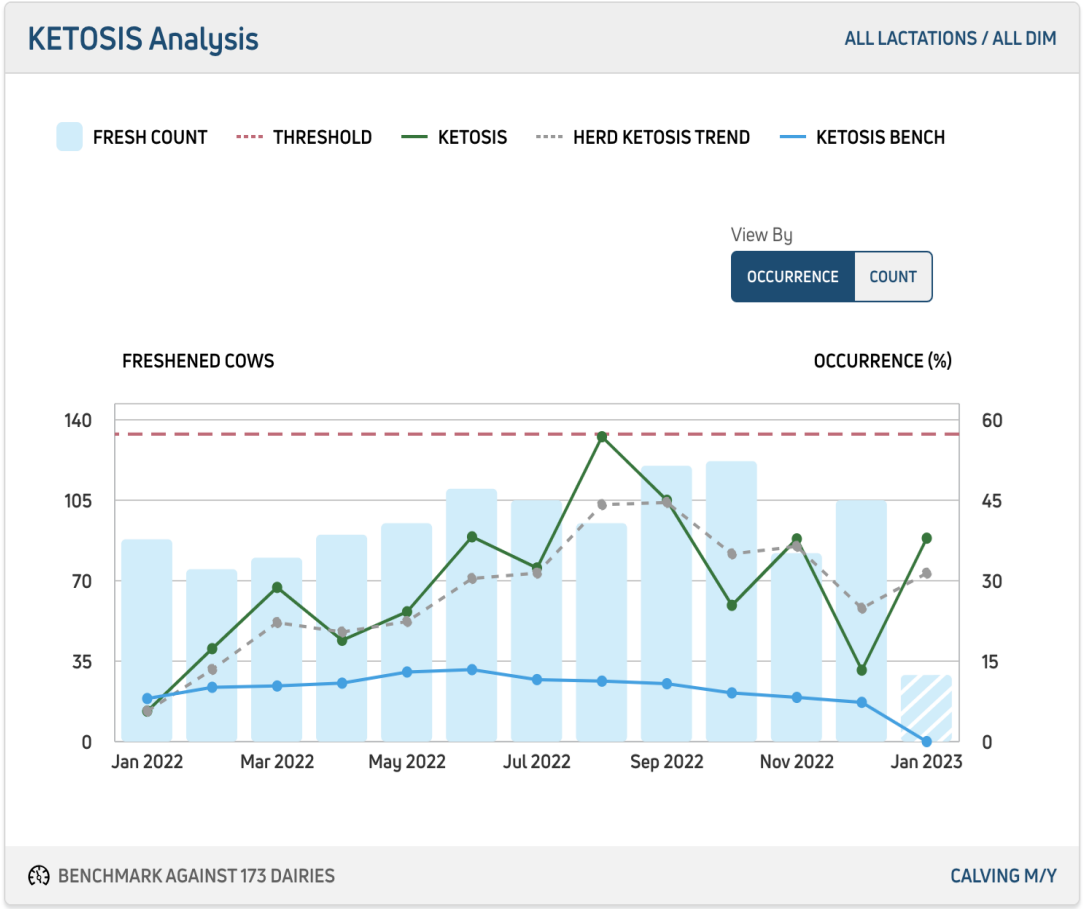
Weaknesses

- User customizability
- Aged desktop software
- Poor analysis options
- Proprietary software



Motivation

- Health data is locked in DC
- Limited integrated analysis
  - Trend lines
  - Benchmarking
- Extensive preprocessing



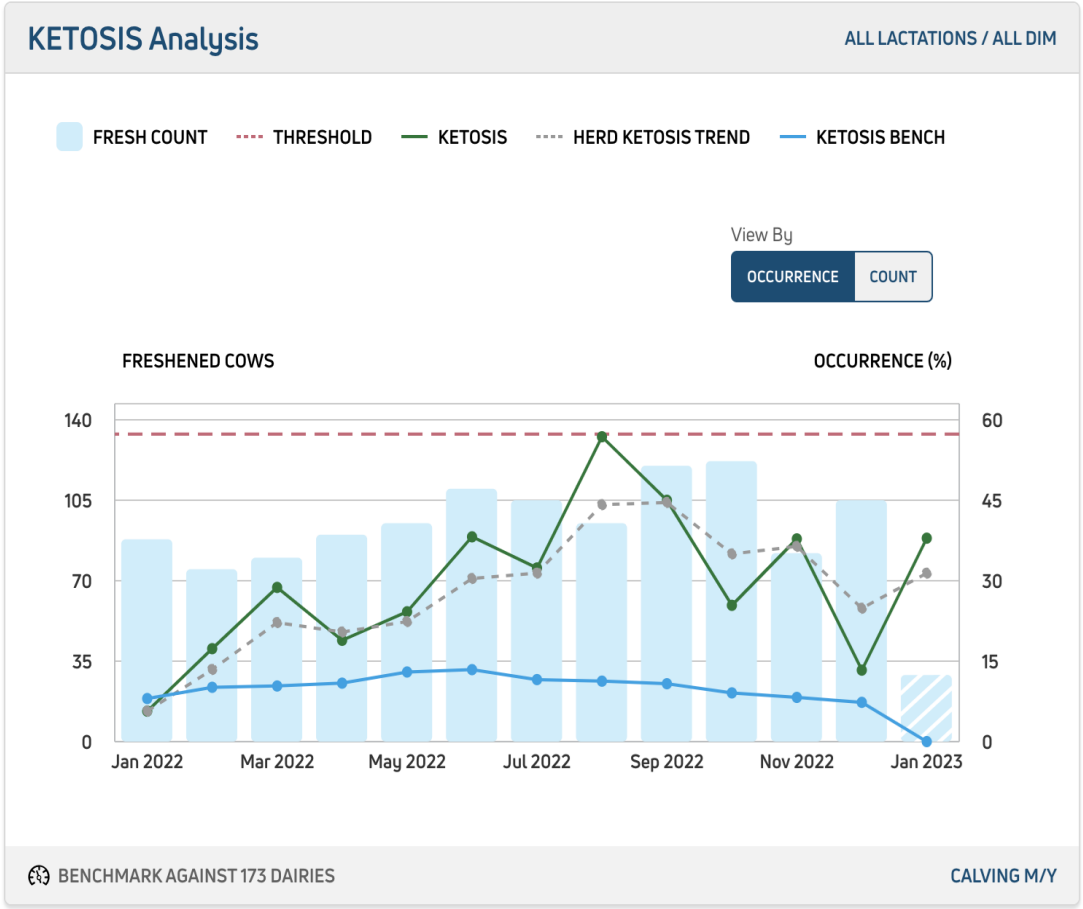
Automate DairyComp

Aim

- Automate DairyComp write out
- Apply tools to support data-driven decisions
- Customize reports and graphs for a better management experience

Audience

- DairyComp users
- Dairy Skills I instructors & students

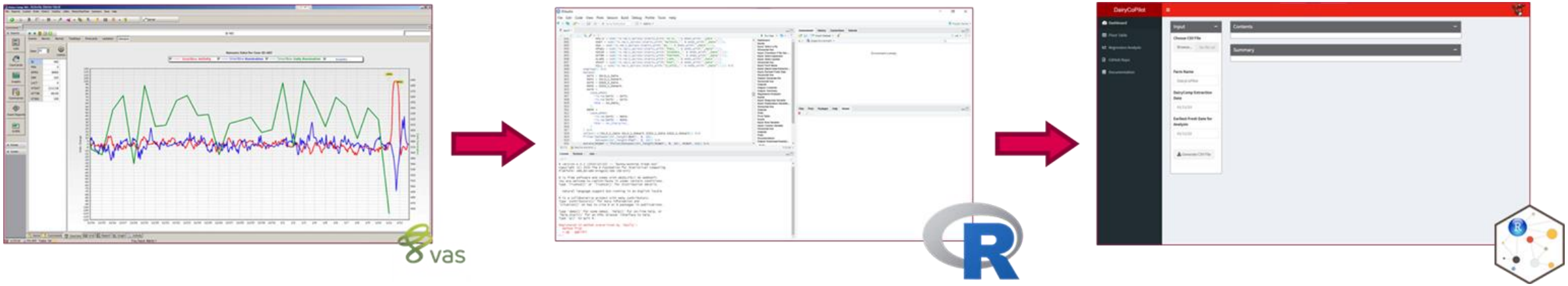


DairyCoPilot

- Simplify data wrangling in DairyComp
  - Time spent creating events < 20 mins
  - Repeatability
- Reduce time for data cleaning
  - Download → Insight < 10 mins
- Simplify analysis
  - GUI
  - Interactive
  - Reportability
- Enhanced analysis
  - Multiple odds ratios → Multiple linear regression

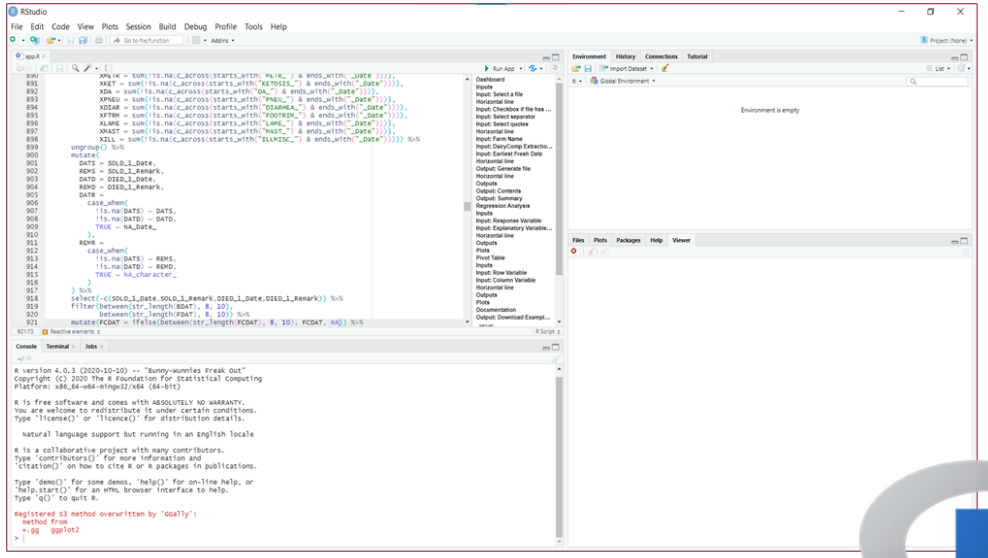
Methods

1. DairyComp Data Extraction
2. DairyCoPilot Data Cleaning
3. Rstudio Connect Deployment



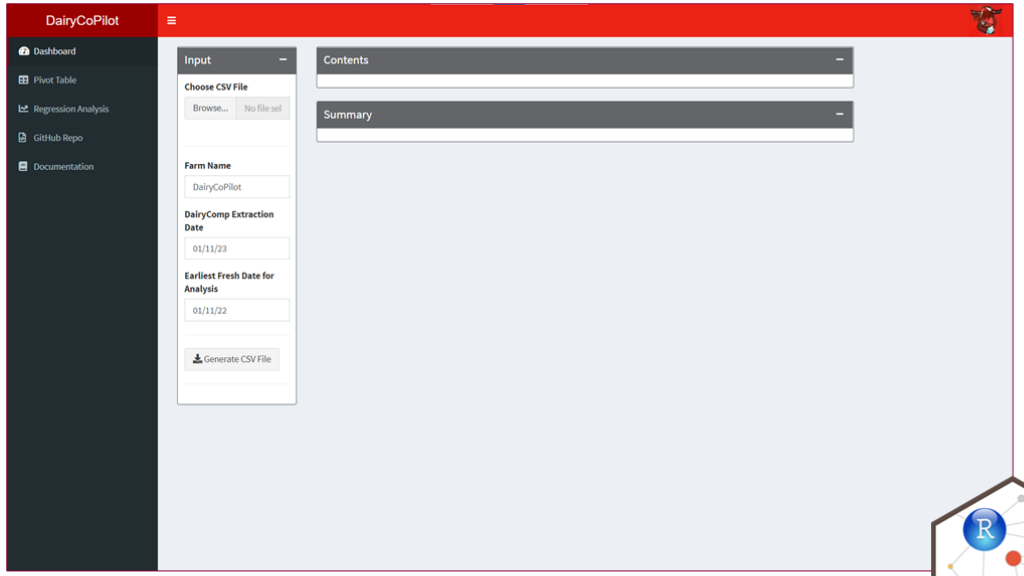
DairyCoPilot Data Cleaning

- Read CSV file
  - Rename events
  - Filter out missing events
  - Create missing columns
- Delete duplicate event records
- Create predefined variables
  - Number of occurrences of an event
  - Date and Remark of first event
  - Cutoff days post calving
- Convert to appropriate variable type



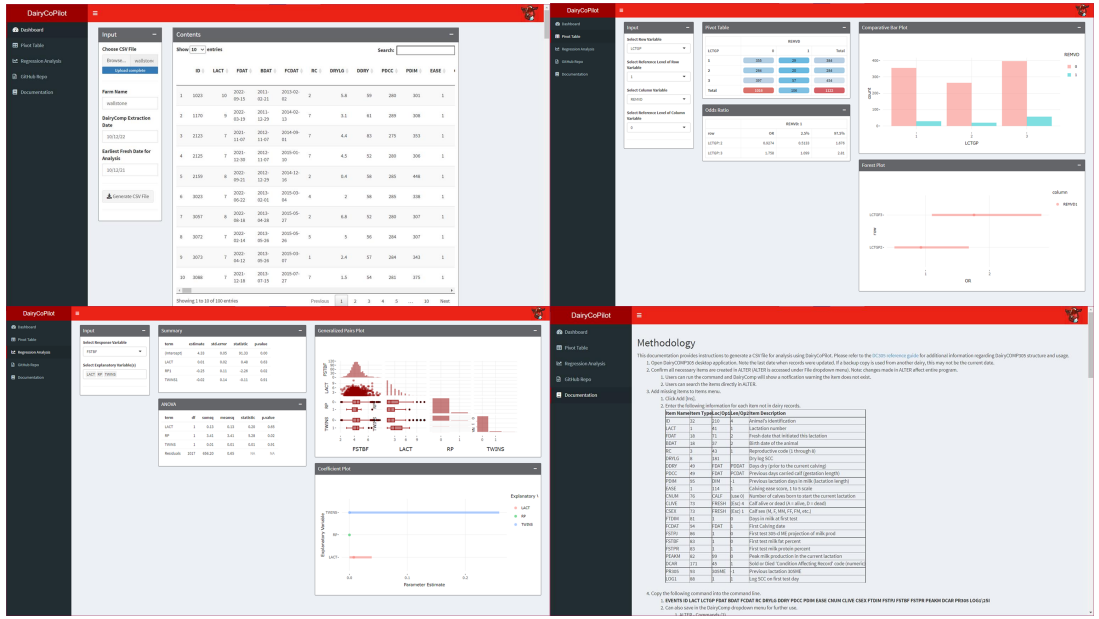
RStudio Connect Deployment

- Standalone publishing platform
  - Dashboards
  - On-demand and scheduled reports
  - Interactive applications
- User-friendly web application
  - Access controls
  - Performance settings
  - View logs



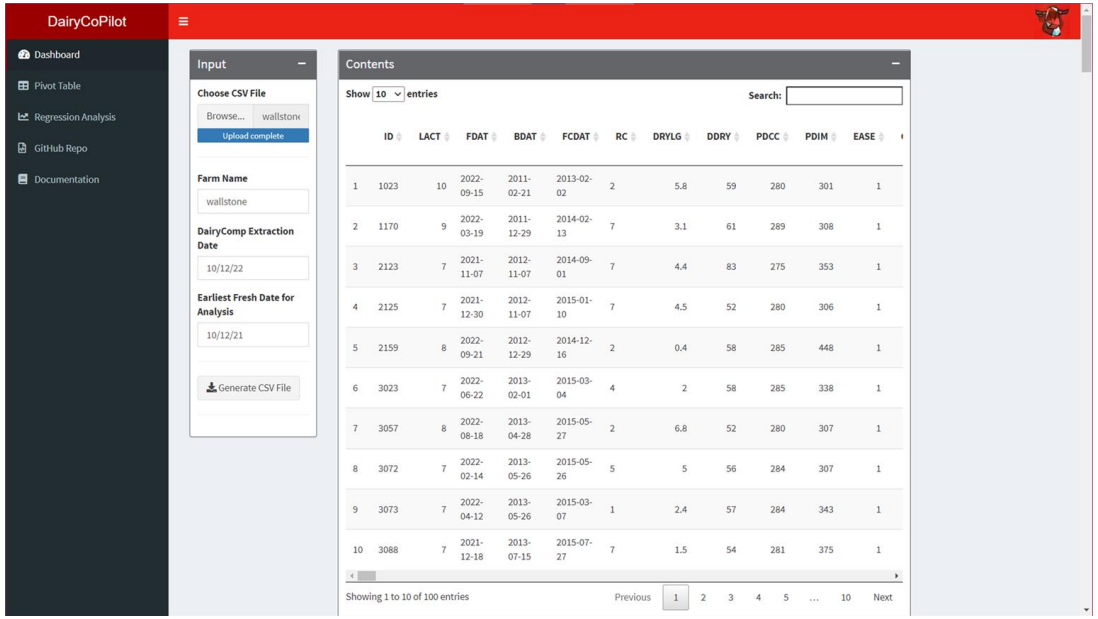
Results

1. Input Tab
2. Pivot Table Tab
3. Regression Analysis Tab
4. Documentation Tab



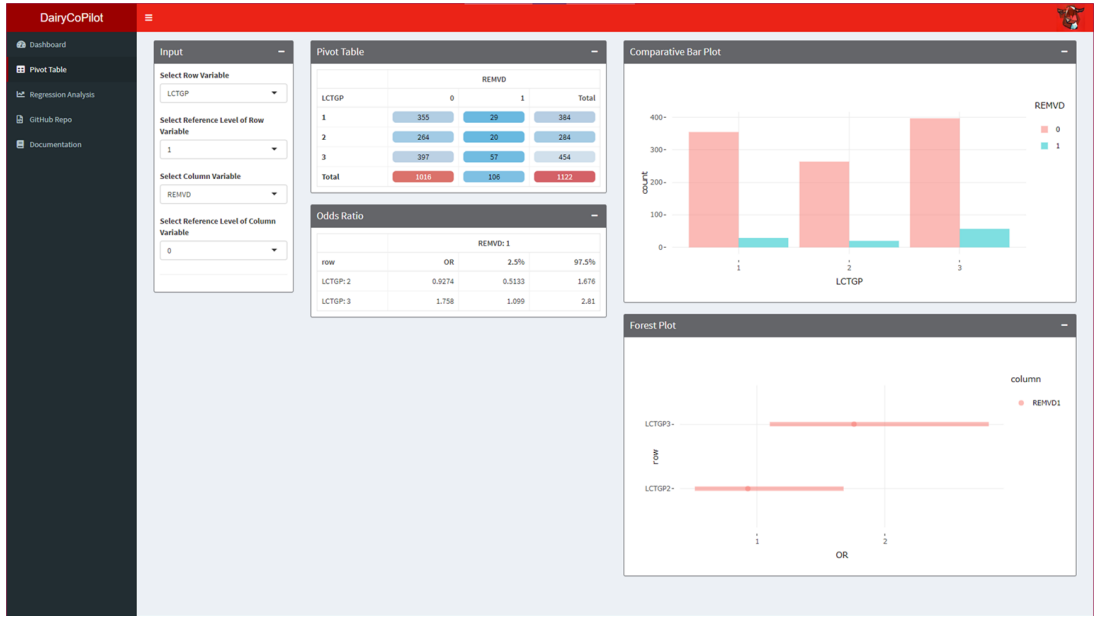
Input Tab

- Upload CSV file from DairyComp
- Download cleaned dataset to local drive
- Descriptive statistics
  - Show top entries
  - Data summary by variable type



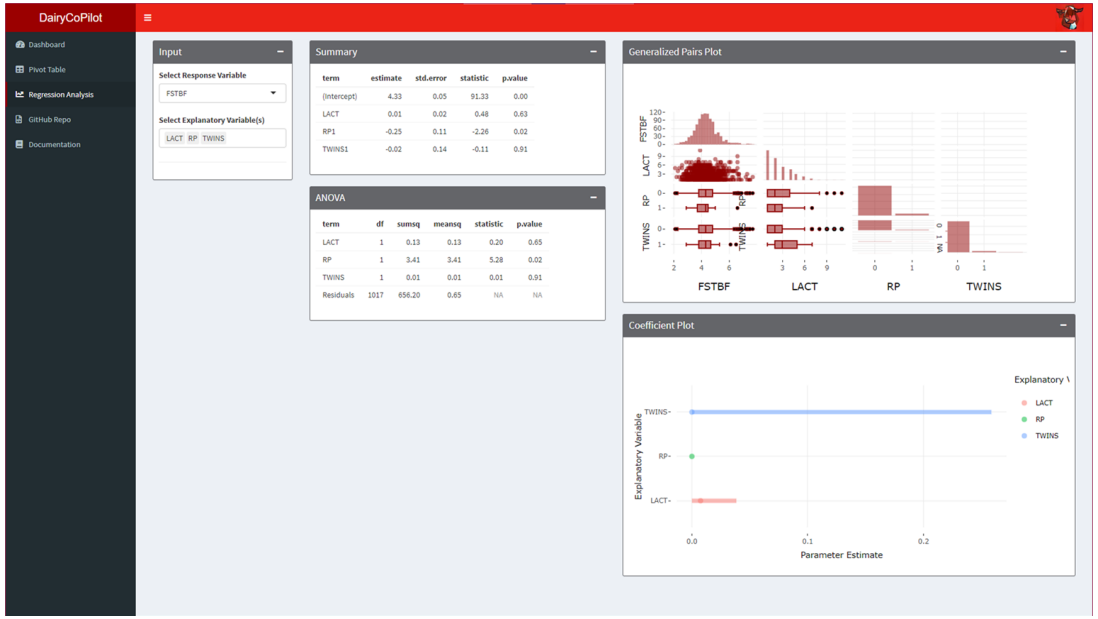
Pivot Table Tab

- Categorical data analysis
  - Select row and column variable
  - Select reference level for each variable
- Statistical analysis
  - Pivot table
  - Odds ratios and 95% confidence intervals
- Data visualization
  - Comparative bar plot
  - Forest plot



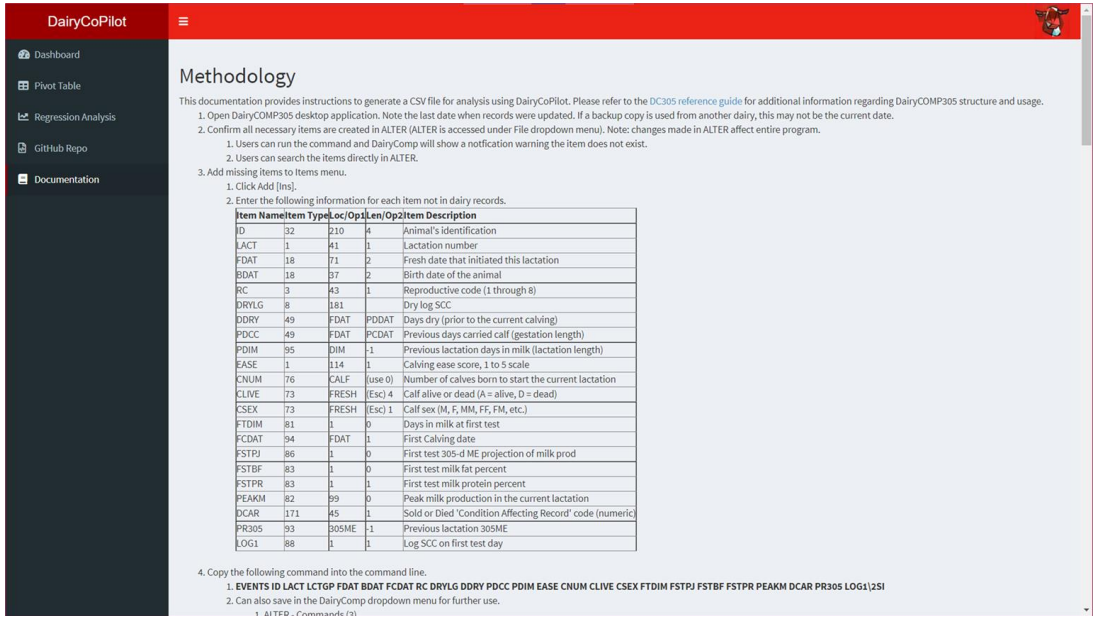
Regression Analysis Tab

- Linear and logistic regression analysis
  - Select a response variable
  - Select multiple explanatory variables
- Statistical analysis
  - Summary table
  - ANOVA table
- Data visualization
  - Generalized pairs plot
  - Coefficient plot



Documentation Tab

- Methodology
  - DairyComp Data Extraction
- Glossary
  - DairyCoPilot Data Cleaning
- Contacts
- Examples
  - Example Datasets
  - Instructional Videos (Upcoming)



Case Study

- Wall-Stone dataset
  - 1000-cow dairy in Wisconsin, USA
  - 43012 events collected in 2022

https://drive.google.com/file/d/11QAM1TTe8YizGLIshxlQsKlFCCOrfjSI/view?usp=sharing

https://data-viz.it.wisc.edu/dairy-copilot/

Summary and Future Directions

- DairyComp is still the backbone of dairy cattle management
- Tools such as DairyCoPilot expand limited analysis options
  - Increase available data
  - Decrease preprocessing time
- Upcoming features
  - Create custom columns
  - Edit factor variables
  - Edit incorrect data directly in application
- Build new features requested by users

SCAN ME

Thank you!