ASBC Annual Meeting 2017

June 4–7 ■ Fort Myers, Florida

See what SCIENCE can brew for you

Simplify QAQC Analyses and Decision Making with Open Source Software

Drew Russey, Ph.D

Laboratory Manager

drew_russey@saintarnold.com

Saint Arnold® Brewing Company

Outline

- What is R?
- R in Brewery QAQC
- Case study: Brewhouse optimization
- Future Directions

The R Project for Statistical Computing

https://www.r-project.org

Pros

- Open source statistical software (FREE)
- Provides powerful, repeatable analyses
- Supplemental packages
- Efficient scripts are reusable and shareable tools
- Auxiliary open source software to improve user interface

Cons

- Steep learning curve
- Time investment

R Base Program

https://www.r-project.org

- Simple functions

```
mean(),   sd(),  sum()
```

- Complex operations e.g.
  - Mean for all beer x

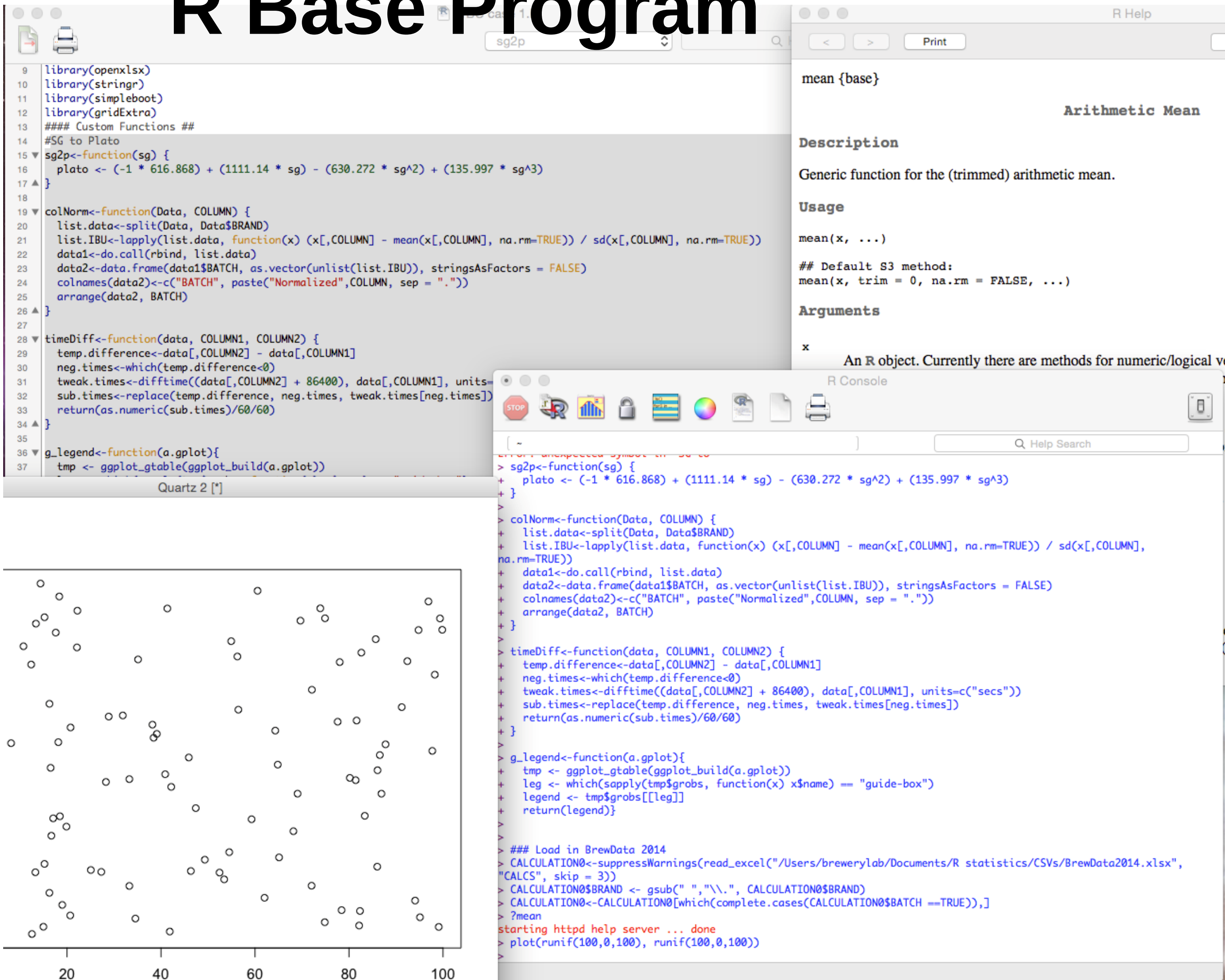
```
lapply(data.as.list, function(x) mean(x$interest))
```

  - Match datasets by batch
    - Merge data output from various equipment

```
data1[match(data2$batch, data1$batch)]
```

R Base Program

https://www.r-project.org



Useful Auxiliary Software

https://www.r-project.org

R: Base program

https://www.rstudio.com

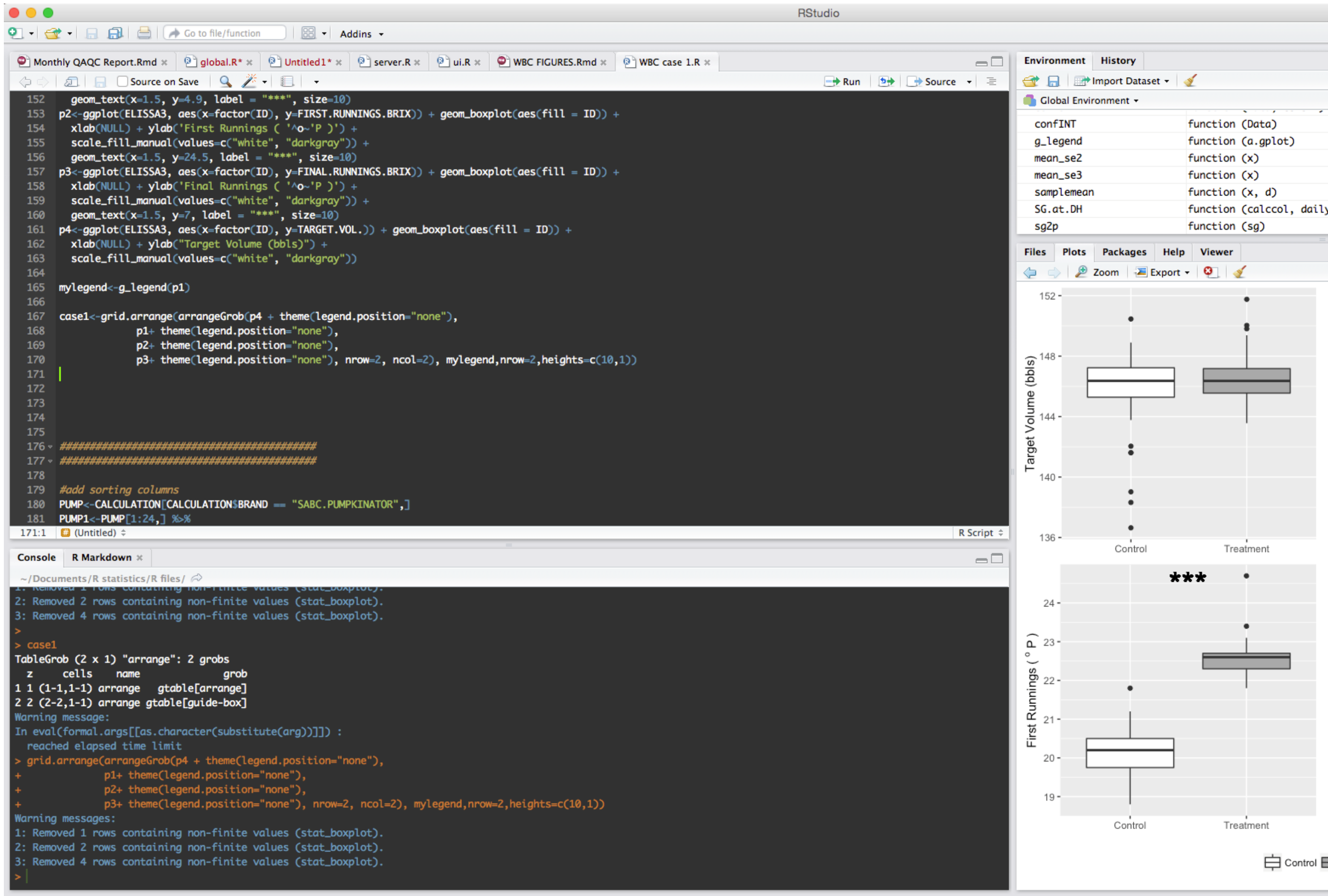
Rstudio: Aids workspace management

- one window instead of 2-4
- debugging support

Useful Auxiliary Software

RStudio

https://www.rstudio.com



Useful Auxiliary Software

https://www.r-project.org

R: Base program

https://www.rstudio.com

Rstudio: Aids workspace management

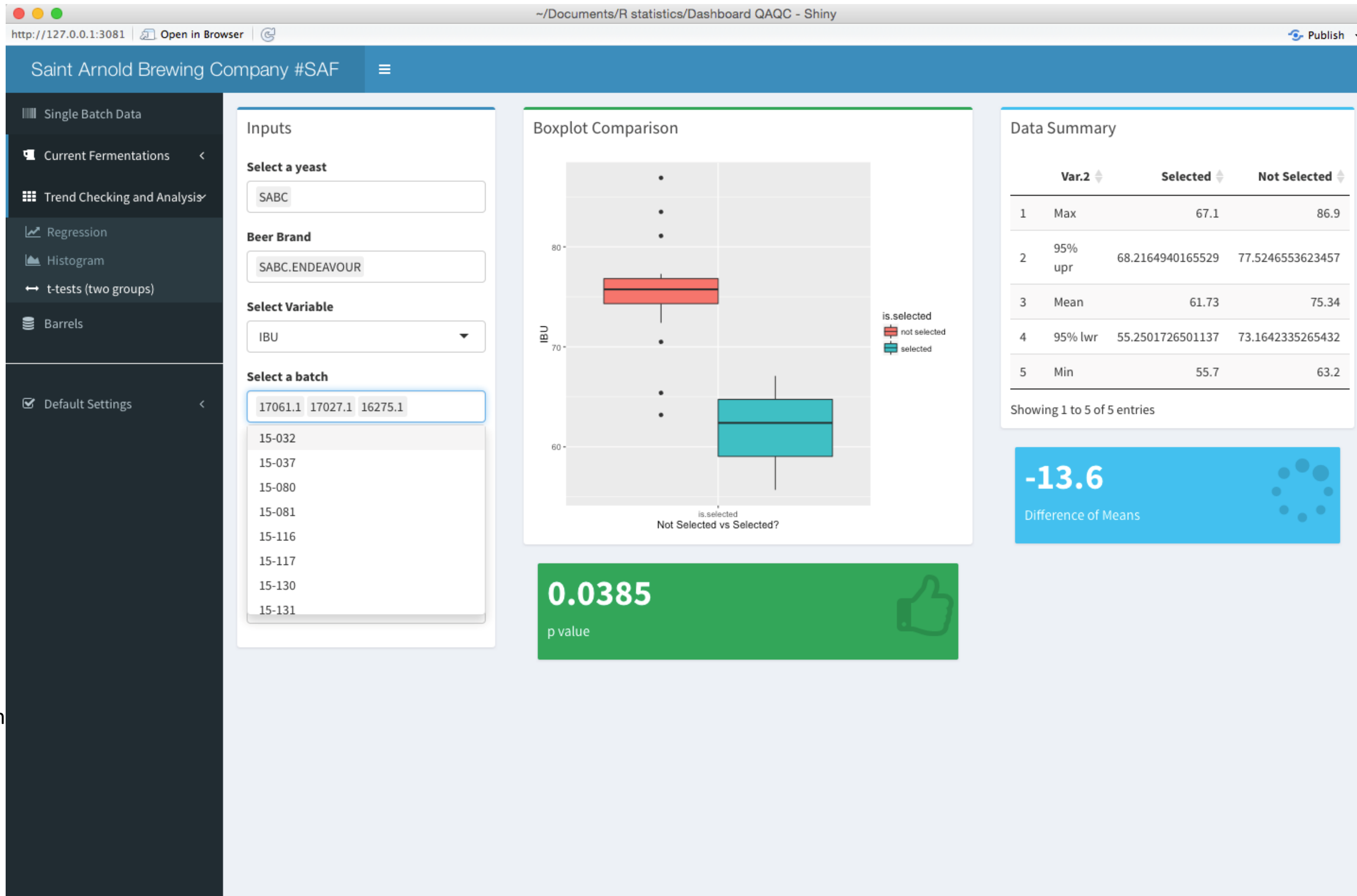
- one window instead of 2-4
- debugging support

http://shiny.rstudio.com

Rshiny: build web applications using only R code. When finished…

- Can perform analyses with pull down menus
- Others can use without ever seeing code.

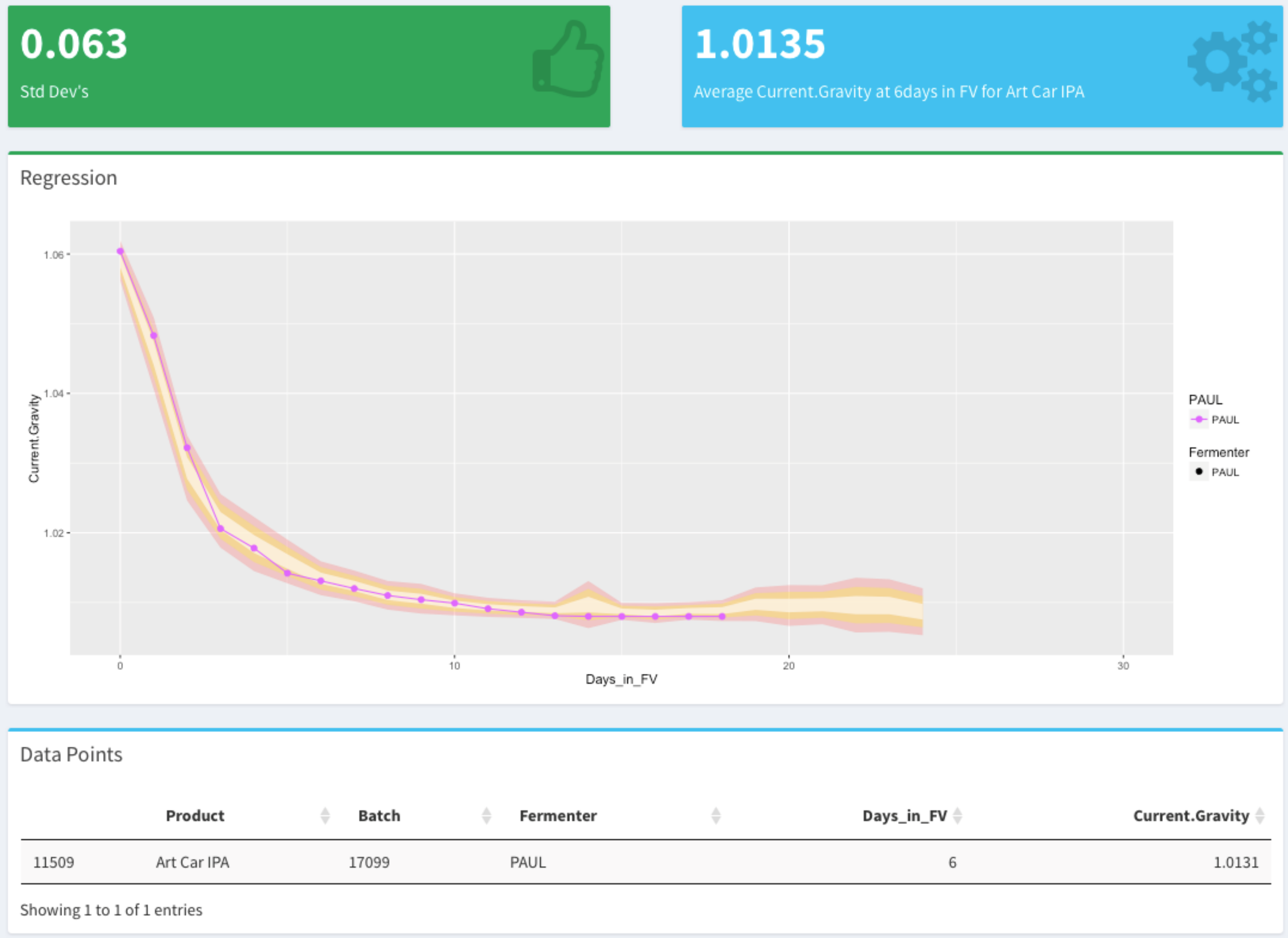
Useful Auxiliary Software



shiny.rstudio.com

R Use Examples

- Track active fermentations
- Check short term and long term trends
- Generate monthly report pdf's
- **Analyze experiments**



Experimental Design

- Step 1: Observation
- Step 2: Form Hypothesis
- Step 3: Plan experiment
  - Ensure results are quantifiable
- Step 4: Power Analysis
  - Aids planning process changes
  - Know when to expect results (determine *n*)
- Step 5: Execute *n* trials
- Step 6: Statistically analyze results of *n* trials
  - *p*-value will provide unbiased answer (significance)
  - *d* will indicate the magnitude of change

R in the Lab

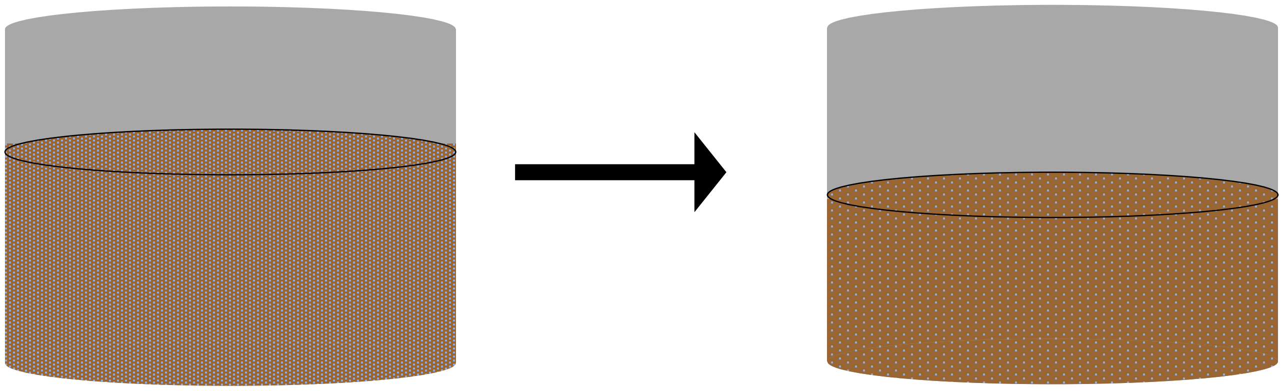
- Step 1: Observation
- Step 2: Form Hypothesis
- Step 3: Plan experiment
  - Ensure results are quantifiable
- Step 4: Power Analysis
  - Aids planning process changes
  - Know when to expect results (determine *n*)
- Step 5: Execute *n* trials
- Step 6: Statistically analyze results of *n* trials
  - *p*-value will provide unbiased answer
  - *d* will indicate the magnitude of change

Brewhouse Optimization: Observation

- Step 1: **Observation**
  - Lauter tun designed to make low Plato beers. Inefficient with higher gravities (IPA).
  - Final runnings were around 5^{o}P

Brewhouse Optimization: Hypothesis

- Step 2: **Hypothesis** Use less grain and thicken mash by reducing mash in water
- Adjusting allocation of water from mash to sparge to more efficiently extract



Brewhouse Optimization: Plan Method

- Step 3: **Plan experiment**
  - Liquor to grist ratio (liquor reduced 6%)
  - Reduce grain bill ~5% to accommodate higher efficiency
  - Drain 14% more wort prior to sparge
  - Increase sparge liquor volume (6%)
  - Calculate Brewhouse Efficiency (Wort Extract/Potential Extract)

For full details: Gutierrez and Russey, 2017, Improving Brewhouse Efficiency by Adjusting Mash Thickness and Lauter and Sparge Volumes

Brewhouse Optimization with R

- Step 4: **Power Analysis**
  - **Power = probability of finding a real effect**
  - `pwr` package in R will solve for n_2 `pwr.t2n.test()`
  - Requires 4 of the 5 below, will solve for remaining variable
    - $\mathbf{n_1}$ = Control group
    - $\mathbf{n_2}$ = Treatment group
    - ***d*** = effect size
    - ***P*** = power. Typically 0.8
    - **α** = Significance level. Typically 0.05

```
pwr.t2n.test(n1=40, p=0.8, sig.level=0.05, …)
```

Brewhouse Optimization with R

- Step 4: **Power Analysis**
  - **Power = probability of finding a real effect**
  - `pwr` package in R will solve for n_2 *`pwr.t2n.test()`*
  - Requires 4 of the 5 below, will solve for remaining variable
    - $\mathbf{n_1}$ = Control group
    - $\mathbf{n_2}$ = Treatment group
    - ***d*** = effect size
    - ***P*** = power. Typically 0.8
    - **α** = Significance level. Typically 0.05

```
pwr.t2n.test(n1=40, p=0.8, sig.level=0.05, …)
```

Effect Size (*d*)

- Many ways to calculate, but essentially…

  the shift in means as standard deviations

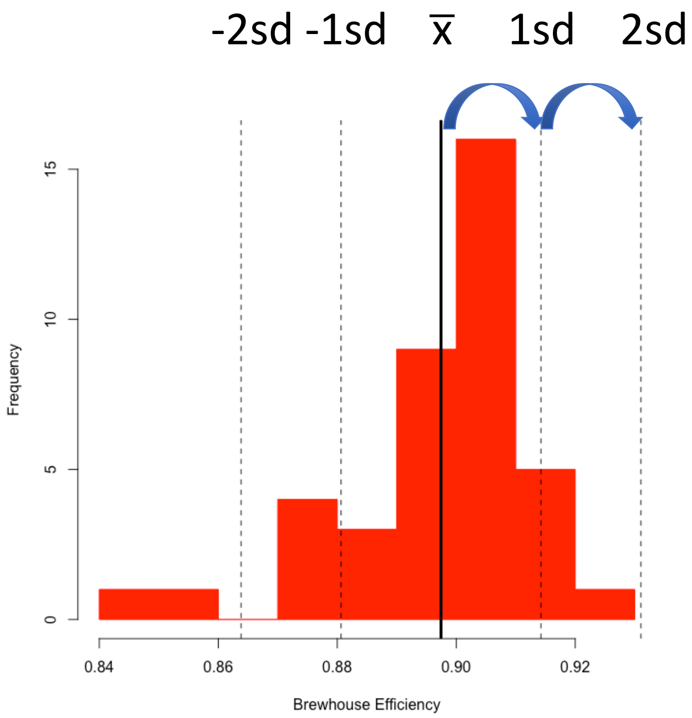
$$d = \frac{\mu_{Treatment} - \mu_{control}}{\sigma_{Control}}$$

```
(mean(treatment) – mean(control)) / sd(control)
```

Effect Size (*d*)

- Many ways to calculate, but essentially…

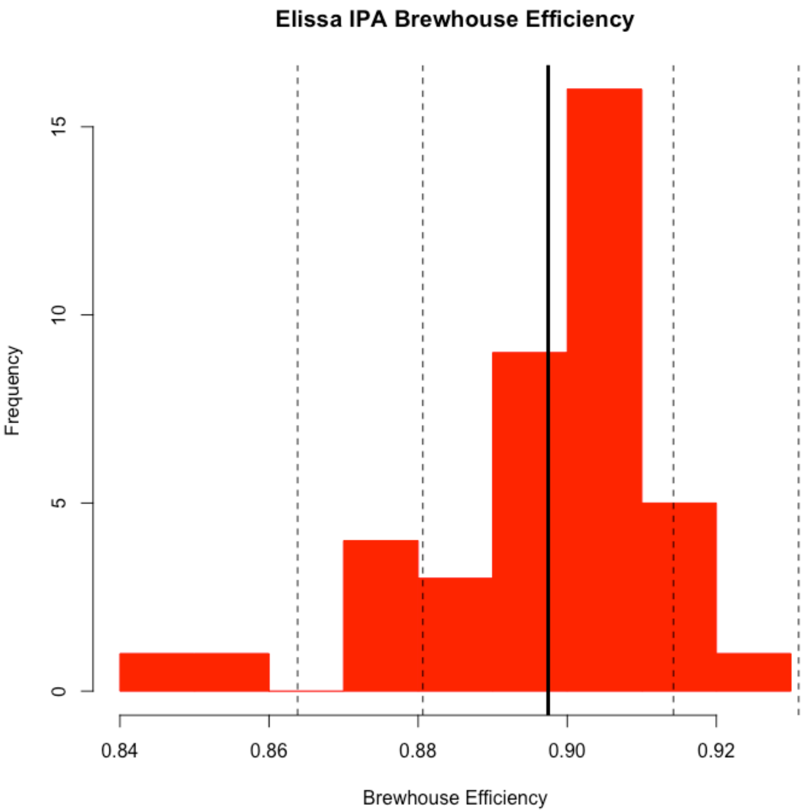
  the shift in means as standard deviations



$$d = \frac{\mu_{Treatment} - \mu_{Control}}{\sigma_{Control}}$$

A.K.A. How many standard deviations will new treatment batches be from controls

Brewhouse Optimization with R



- Step 4: Power Analysis

Determine when to expect results (determine needed *n*)

  - Variables for `pwr.t2n.test()`

n_1 = 40 mashes from previous year

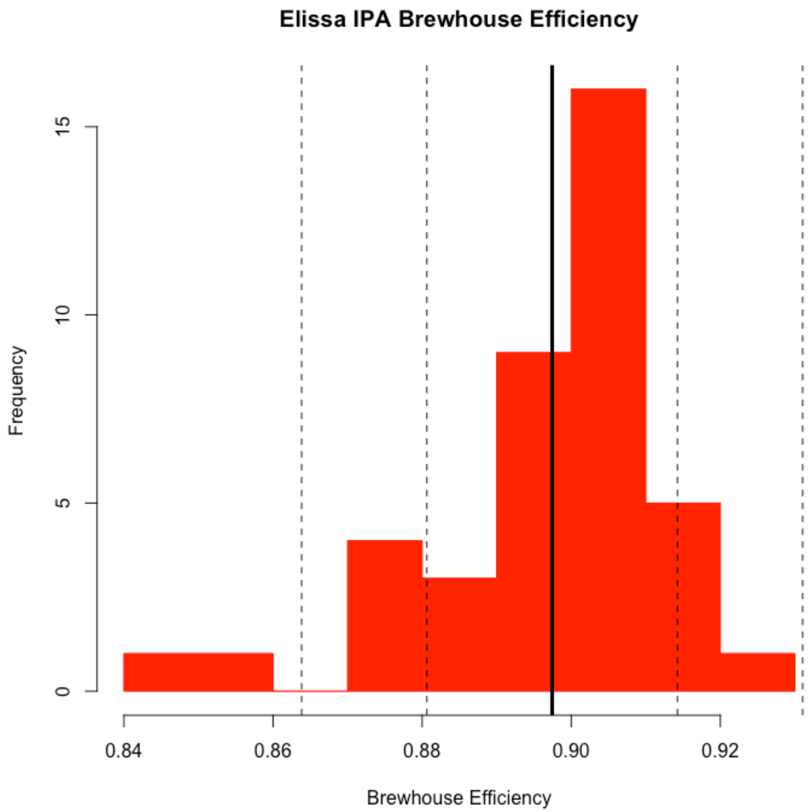
P = 0.8

α = 0.05

d =

Description	d	n2 needed
Small	0.3	∞
Medium	0.5	153
Large	0.8	19
Huge	2.0	3

Brewhouse Optimization with R



- Step 4: Power Analysis

Determine how large of an effect you could find

- Variables for `pwr.t2n.test()`
  - n_1 = 40 mashes from previous year
  - P = 0.8
  - α = 0.05
  - n_2 =

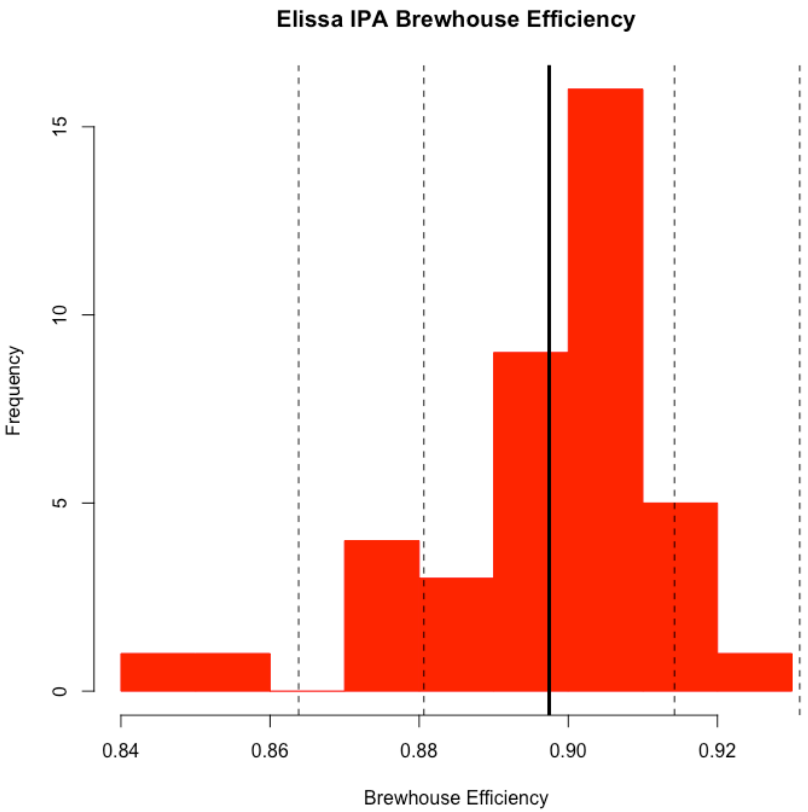
n_2	Min *d*	Description
3	1.71	Large
6	1.25	Large
10	1.01	Large
34*	0.66	Medium

Brewhouse Optimization with R

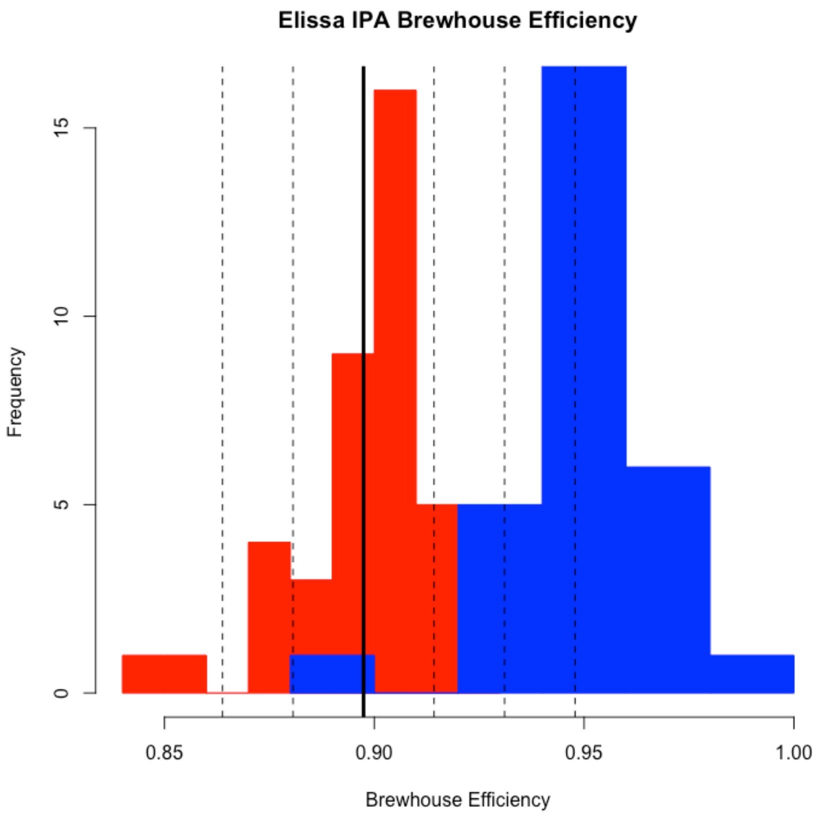
- Step 1: Lauter tun designed to make low Plato beers. Inefficient with higher gravities such as IPA.
- Step 2: Adjusting allocation of water from mash to sparge will enhance brewhouse efficiency
- Step 3: Reallocate water from mash to sparge, condense grist, quantify brewhouse efficiency
- Step 4: Power Analysis
  - Determine number of trials necessary and/or
  - Determine how large an effect needed given n trials
- **Step 5: Execute**
- Step 6: Analyze results with t-test equivalent (Mann-Whitney u-test)

Increase in Brewhouse Efficiency

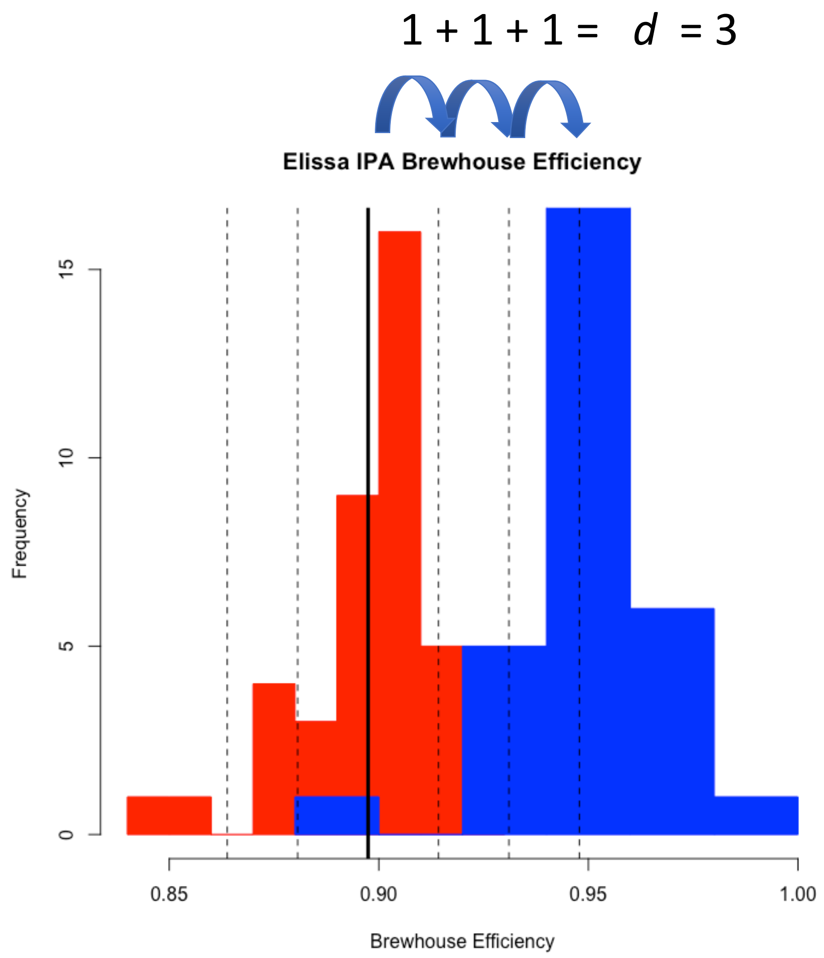
-2sd -1sd x 1sd 2sd



89.7% 94.9%



Increase in Brewhouse Efficiency



$$d = \frac{\mu_{Treatment} - \mu_{Control}}{\sigma_{Control}}$$

$$d = \frac{94.9\% - 89.7\%}{1.68\%}$$

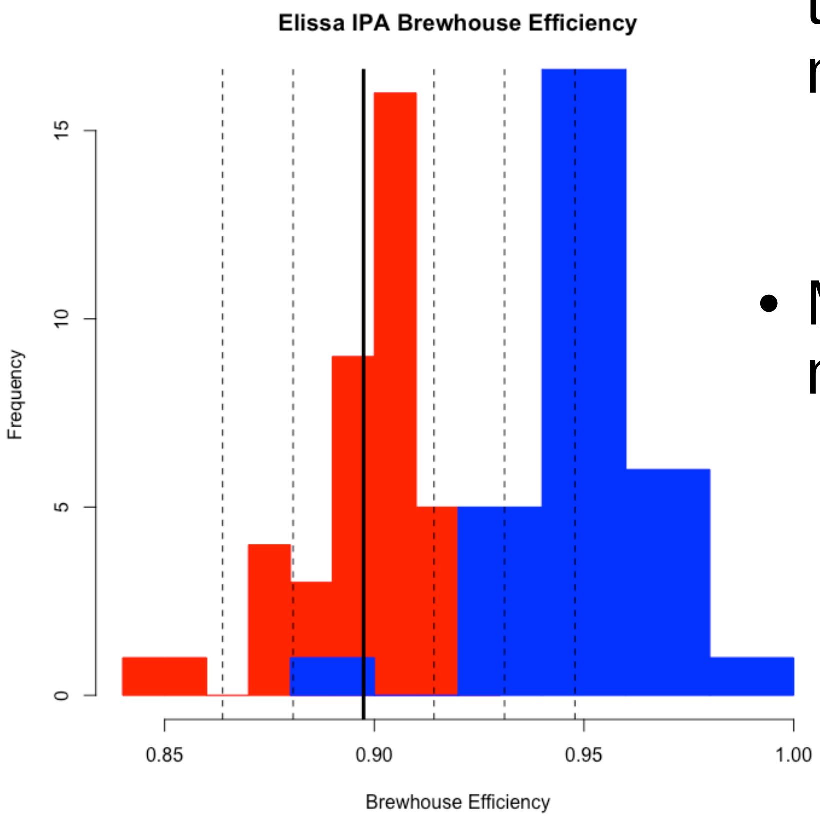
d = 3.09 a “huge” effect

But is it significant?

Brewhouse Optimization with R

- Step 1: Lauter tun designed to make low Plato beers. Inefficient with higher gravities such as IPA.
- Step 2: Adjusting allocation of water from mash to sparge will enhance brewhouse efficiency
- Step 3: Reallocate water from mash to sparge, condense grist, quantify brewhouse efficiency
- Step 4: Power Analysis
  - Determine number of trials necessary and/or
  - Determine how large an effect needed given n trials
- Step 5: Execute
- **Step 6: Analyze results with t-test equivalent (Mann-Whitney u-test)**

Increase in Brewhouse Efficiency



- Shapiro test on both control and treatment groups to test for normality
  - `shapiro.test(efficiency)`
  - $p = < 0.05$, non-normal
- Mann-Whitney/Wilcoxon test is a non-parametric t-test equivalent
  - Safe to use with non-normal distribution
  - `wilcox.test(efficiency_red, efficiency_blue)`
  - $p = 2.45 \times 10^{-12}$

https://www.r-project.org

FutuRe DiRections

- Increase awareness of open source software for brewing and brewery lab analyses
- Create a community of R users to share targeted scripts
  - Make analyses accessible to breweries of all sizes
  - Standardize methods of analysis alongside measurement
  - Increase reproducibility and transparency of beer science
- All base R scripts from this presentation available

Questions?

SAINT ARNOLD BREWING COMPANY

drew_russey@saintarnold.com

Special Thanks to

Eddie Gutierrez
Saint Arnold Brewing Company
eddie_guttierez@saintarnold.com