# Black and Gold Analytics presents: The BUSINESS **ANALYTICS** **BATTLE**

## Introduction

Lealand Morin, Joushua Eubanks

Department of Economics
College of Business Administration
University of Central Florida

March 30, 2023

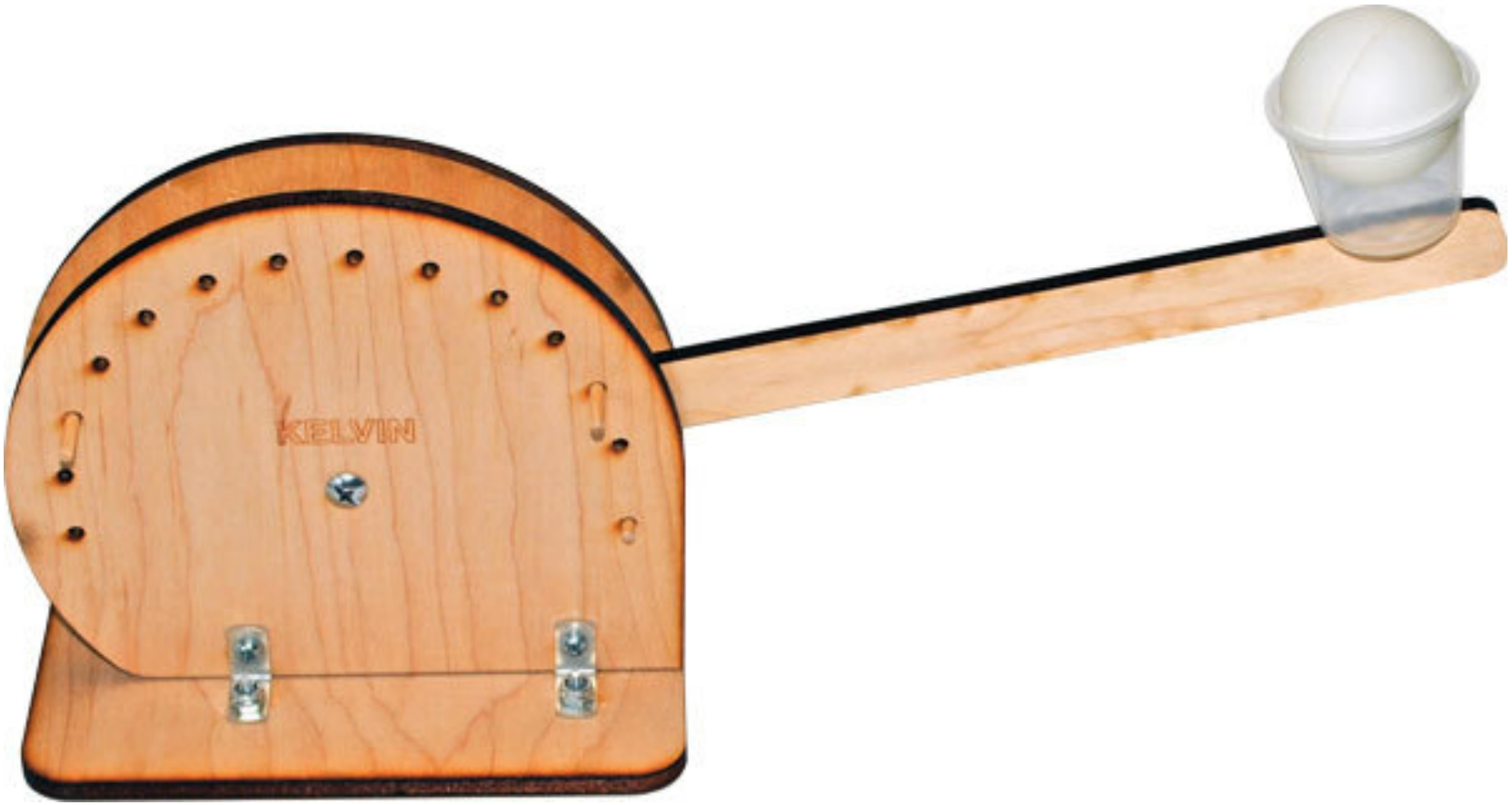
# An Exercise in Predicting Continuous Variables

*Definition* **(Continuous Variables)**: A variable with an uncountable set of outcomes
(That is, the set of outcomes is typically limited only by the precision of measurement)

- Outcomes measured in dollars
  - Sales (and returns), reported income
- Outcomes measured as rates
  - Conversion rates, click-through rates
- Outcomes with a large number of outcomes
  - FICO scores, number of drivers, houses or building permits

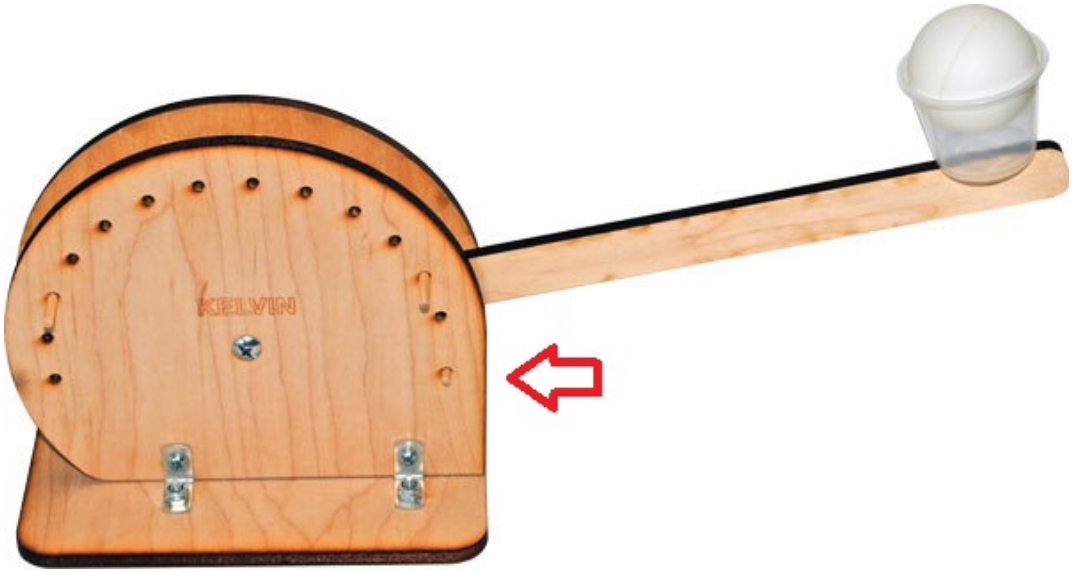
# The Statapult

Collect data by conducting an experiment



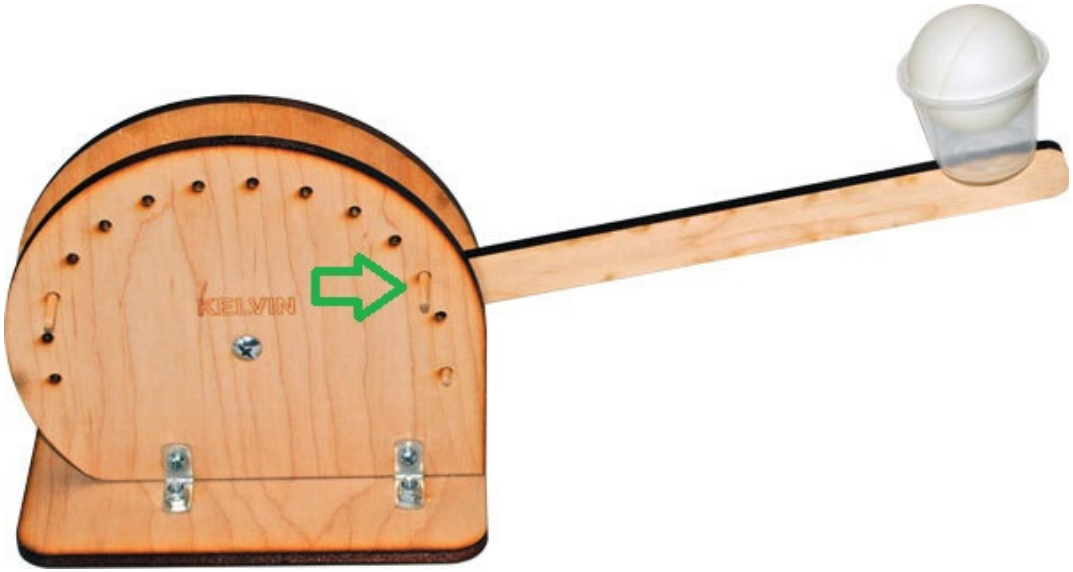
# The Statapult: Settings

Choose a starting point to set spring tension



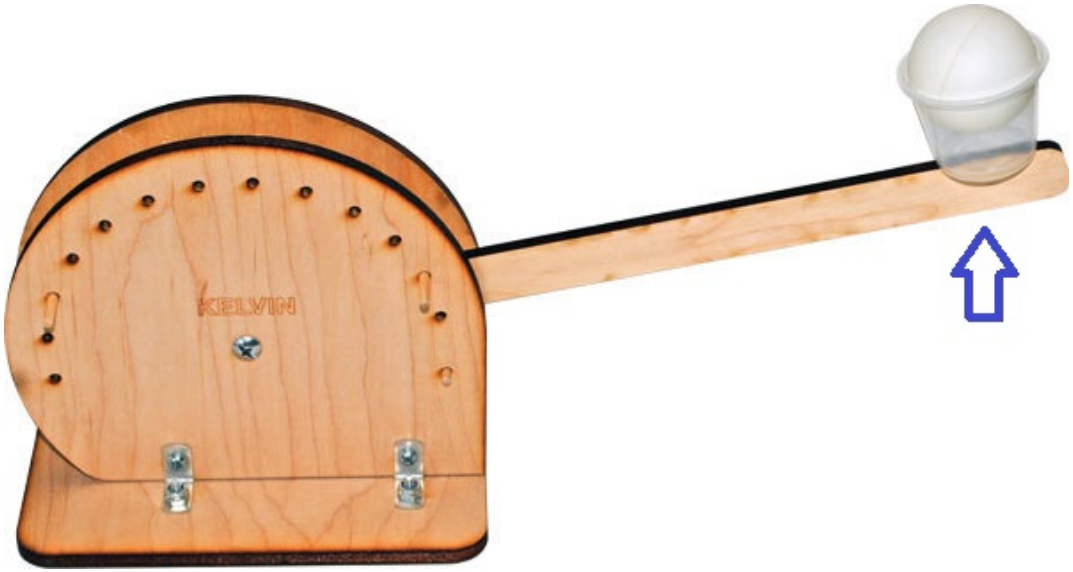
# The Statapult: Settings

Choose a stopping point to set launch angle



# The Statapult: Settings

Choose a cup position to set lever length



# Data Collection

Conduct your experiment

- Set the catapult on one end of the black sheet
- Align the measuring tape with the end of the catapult
- For each launch, record the distance the ball traveled:
  - Roll the ball in baby powder
  - Launch it with the catapult
  - Record the distance shown by the white spot on the black cloth
- Repeat for all 18 combinations of catapult settings:
  - 3 starting points
  - 3 stopping points
  - 2 cup positions

# Analyze the Data

Build a model for the distance thrown with the catapult

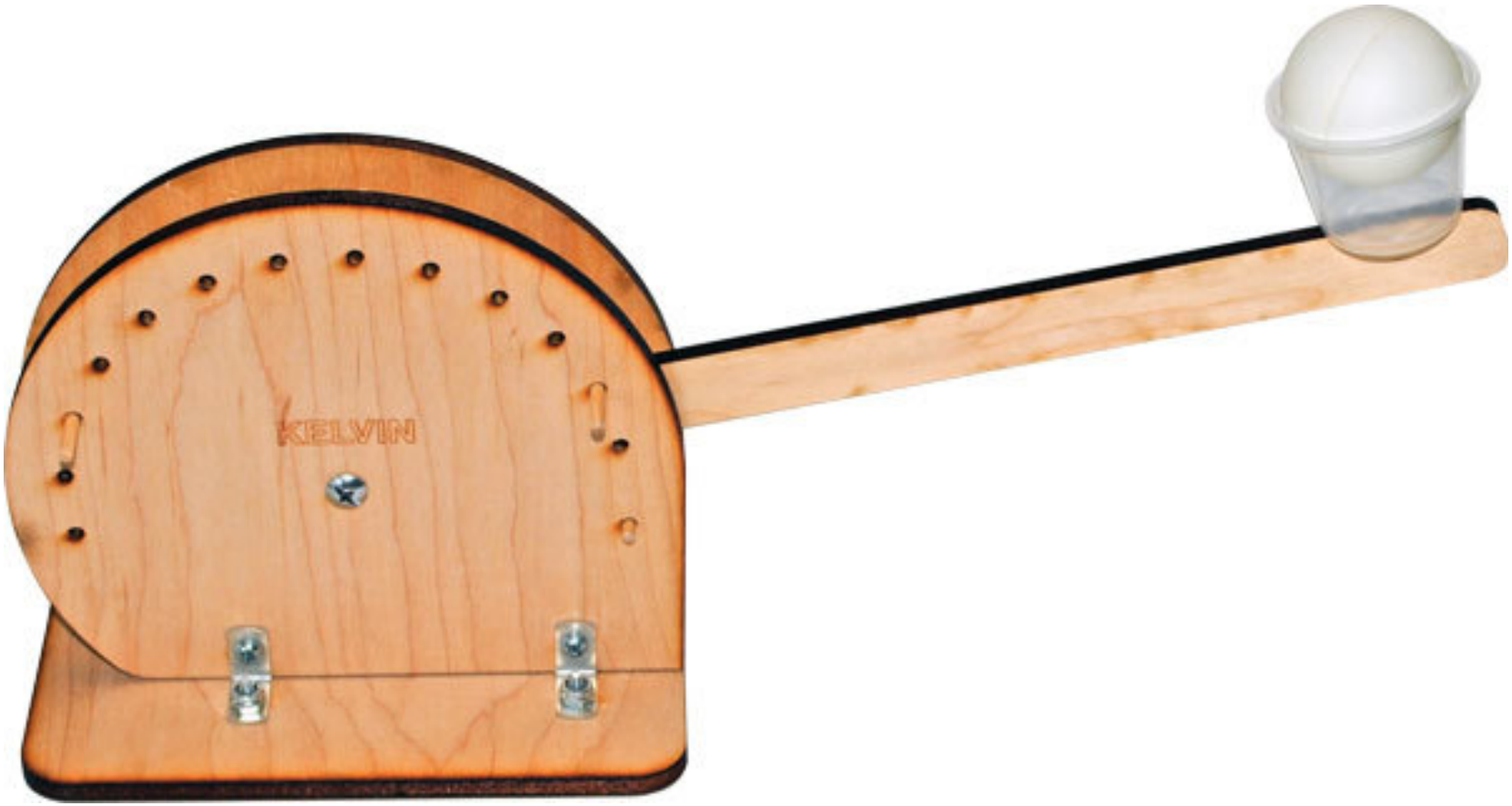
```
Call:
lm(formula = house_price ~ prediction + in_cali + earthquake,
    data = housing_data[1:num_obs_estn, ])

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.16828    0.16665  -1.010    0.315
pred[illegible]  1[illegible]  27[illegible]  5[illegible]  1[illegible]e-05
in_ca[illegible] -0[illegible]  0[illegible]  .818
earth[illegible]ak  -[illegible]  0[illegible]  -  346  <[illegible]e-16
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1004 on 96 degrees of freedom
Multiple R-squared:  0.7129,    Adjusted R-squared:  0.704
F-statistic: 79.48 on 3 and 96 DF,  p-value: < 2.2e-16
```

# The Statapult

Test your model by conducting another experiment



# Testing Your Model

Conduct your experiment

- Set the catapult on one end of the black sheet
- Align the measuring tape with the end of the catapult
- **Specify the goal: the target distance to hit with the ball**
- Take turns and record the distance the ball traveled:
  - Roll the ball in baby powder
  - Launch it with the catapult
  - Record the distance shown by the white spot on the black cloth
- Finally, **compare the outcome to the goal**
- The closest to the target, without going over, **wins!**

# Case Study: Response Rates

Zillionz

- Zillionz is a popular search engine
- Their business is based on selling advertisements
- The value for their customers depends on response rates
- They observe the customer's catapult settings
- They choose whether or not to fire the catapult
- The payoff is as follows:
  - Gain \$10 if roll $d > 60$ inches
  - Lose \$1 otherwise

# Case Study: Response Rates

How does this relate to firing the catapult?

**Data Input:**
Customer's settings
S

**Prediction:**
$P\{d|S\}$


**Decision:**
Launch catapult (or not)


**Outcome:**
+\$10 if $d > 60$
else -\$1

# See the Results

Conduct your experiment

- Set up the catapult as for the previous experiments, according to the settings specified by the customer
- **Before each launch, decide whether to accept the launch**
- Record the distance thrown in a separate column, depending on whether it was accepted or rejected
- Repeat for all 10 catapult settings on the worksheet
- Finally, calculate your total score

# Case Study: Forecasting Sales

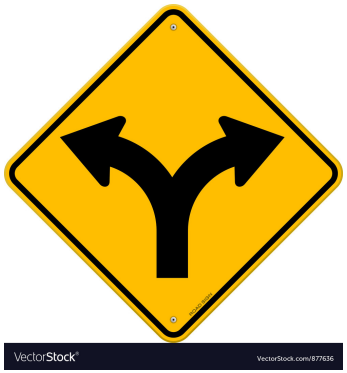

- Fairway is an online retailer
- Their sales fluctuate throughout the year
- To manage their inventory, they need to forecast sales
- They observe the catapult settings
- They choose an inventory level $\hat{d}$
- They observe the distance thrown $d$
- The payoff is as follows:
  - If they underestimate distance: $\hat{d} < d$, lose \$10 for each inch under
  - If they overestimate distance: $\hat{d} > d$, lose \$5 for each inch over

# Case Study: Forecasting Sales

How does this relate to firing the catapult?

**Data Input:**
Catapult settings
S

**Prediction:**
$P\{d|S\}$




**Decision:**
Purchase inventory
$\hat{d}$

**Outcome:**
-\$10 if too high
else: -\$5 if too low

# See the Results

Conduct your experiment

- Set up the catapult as for the previous experiments
- Observe the catapult settings
- **Before each launch, predict the distance thrown**
- Launch with the prescribed settings
- Repeat for all 10 catapult settings on the worksheet
- Finally, calculate your total score

# Case Study: Pricing Policy


- Homefix is a retailer of home improvement products
- They sell a wide variety of products
- Their goal is to choose prices to maximize profits
- They observe two of the three catapult settings
- They choose the remaining setting
- The payoff is as follows:
  - Gain \$10 if the ball lands between 54 and 60 inches
  - Lose \$1 for every inch out of range

# Case Study: Pricing Policy

How does this relate to firing the catapult?

**Data Input:**
Catapult settings $S_1$

**Prediction:**
$P\{d|S_1, S_2\}$ for each $S_2$


**Decision:**
Set remaining setting $S_2$


**Outcome:**
Gain \$10 if close
Lose \$1 for distance out of range

# See the Results

Conduct your experiment

- Set up the catapult as for the previous experiments
- Observe the first two catapult settings
- Before each launch, predict the distance thrown with each remaining option
- **Choose an optimal setting for the remaining option**
- Launch with the chosen and prescribed settings
- Repeat for all 10 catapult settings on the worksheet
- Finally, calculate your total score

# Any questions?