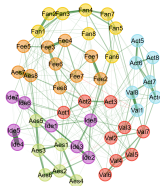
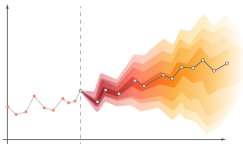


Introduction

DS-5740 Advanced Statistics



Some Introductions

Instructor

Alexander Christensen

Psychology and Human Development

Peabody College

Teaching Assistant

Marisa Blackman

Department of Biostatistics

Vanderbilt School of Medicine

Some Introductions

- 1 Name that you'd like me and other people to use
- 2 Area of study
- 3 Topics you're interested in (e.g., finance, psychology, social media, sports)
- 4 What you'd like to get out of this course
- 5 Career goals (what would you like to do?)





You and your group have just survived the crash of a small plane...

Both the pilot and co-pilot were killed in the crash. It is mid-January, and you are in Northern Canada. The daily temperature is -25°F (or -32°C) and the night time temperature is -40°F (or -40°C). There is snow on the ground, and the countryside is wooded with several creeks criss-crossing the area. The nearest town is 20 miles away. You are all dressed in city clothes appropriate for a business meeting.

Your group of survivors managed to salvage the following items:

- a ball of steel wool
- a small ax
- a loaded .45-caliber pistol
- can of Crisco shortening
- newspapers (one per person)
- cigarette lighter (without fluid)
- extra shirt and pants for each survivor
- 20 × 20 ft. piece of heavy-duty canvas
- a sectional air map made of plastic
- one quart of 100-proof whiskey
- a compass
- family-size chocolate bars (one per person)

Group A

Rank the items according to their importance to your survival, starting with **1** for the most important item to **12** for the least important. Work quickly, alone, and *don't* discuss the problem with anyone in your group. Record your rankings in the scoring table down column **A** labeled **Individual Ranking**.

Group B

As a group, discuss what items are most critical for survival. You should try to come to a consensus as the more everyone agrees, the more likely you and your group will survive. Once there is a consensus, record the group rankings in the scoring table down column **B** labeled **Group Ranking**.

Group A

As a group, discuss what items are most critical for survival. You should try to come to a consensus as the more everyone agrees, the more likely you and your group will survive. Once there is a consensus, record the group rankings in the scoring table down column **B** labeled **Group Ranking**.

Group B

Rank the items according to their importance to your survival, starting with **1** for the most important item to **12** for the least important. Work quickly, alone, and *don't* discuss the problem with anyone in your group. Record your rankings in the scoring table down column **A** labeled **Individual Ranking**.

- | | | |
|---|---|--|
| 1 cigarette lighter (without fluid) | 5 20 × 20 ft. piece of heavy-duty canvas | 9 a loaded .45-caliber pistol |
| 2 a ball of steel wool | 6 a small ax | 10 one quart of 100-proof whiskey |
| 3 extra shirt and pants for each survivor | 7 family-size chocolate bars (one per person) | 11 a compass |
| 4 can of Crisco shortening | 8 newspapers (one per person) | 12 a sectional air map made of plastic |

absolute value

Individual Accuracy : $|A - C|$

Group Accuracy : $|B - C|$

Influence : $|A - B|$

- 0-12: Great survival skills – rescued!
- 13-24: Above average survival skills – rescued!
- 25-36: Frostbit, hungry, and tired – rescued!
- 37-48: Dehydrated and barely alive – it was tough, but rescued!
- 49-60: Rescued, but only just in time! Lost a few toes.
- 61-72: Sorry, only scraps of your clothing were recovered months after the search was called off.

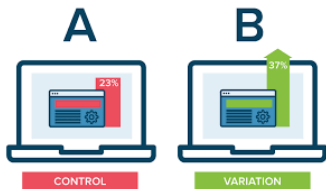
Who did better: you or the group?

- Problem space
- Confirmation bias
- Functional fixedness

- Groupthink
- Social loafing
- Group dominance

Did people in group B perform better *individually* than people in group A?

You'll answer this question as an assignment later in the course



Preparing for the Course

Forecasting the Course

- Forecasting (weeks 1-7)
- Catch-up/Time for First Project (week 8)
- Dimension Reduction/Clustering (weeks 9-12)
- A/B Design or Introduction to Bayesian Statistics (weeks 13-14)
- Catch-up/Time for Second Project (week 15)

Software for the Course



[Download R](#)



[Download Rtools](#) (Windows only)



[Download RStudio](#)

```
install.packages("package_name")
```

- Replace package_name with the name of the package
- For example:

```
install.packages("fpp3")
```

Packages to download

- Forecasting: {fpp2}, {fable}, {prophet}, {tsibble}, {seasonal}, {forecast}, {tidyquant}
- Dimension Reduction/Clustering: {psych}, {cluster}, {factoextra}
- Quasi-experimental Design: {scidesignR}

```
devtools::install_github("user/package_name")
```

- Replace user with the username of package developer
- Replace package_name with the name of the package
- For example:

```
devtools::install_github("hfgolino/EGAnet")
```



Load R Packages

- Load the {fpp3} package

```
library(fpp3)
```

- Possible error

```
Error in library(fpp3): there is no package called  
'fpp3'
```

Got an Error?

Day 1 of Programming



10 Years of Programming



● Google

● StackOverflow

● Cross Validated

Tip: copy and paste error output verbatim

Pick a Stock

Pick a Stock | First Assignment

As a semester long assignment, you will track a stock of your choice. It can be any stock, so be sure to pick one that you're interested in. You cannot choose the same stock as another person – the stock must be your stock and your stock only.

Halfway through the semester, you will visualize the time series of the data using methods taught over the next several weeks and you'll construct a forecasting model for your stock using the last three years of monthly data. You will turn in this model.

At the end of the semester, we will check in on how the models have done. The goal of this assignment is to use the skills you've accumulated through the forecasting component of this course. Your grade will *not* be affected by how well your model forecasts the stock (stocks are **hard** to forecast!).

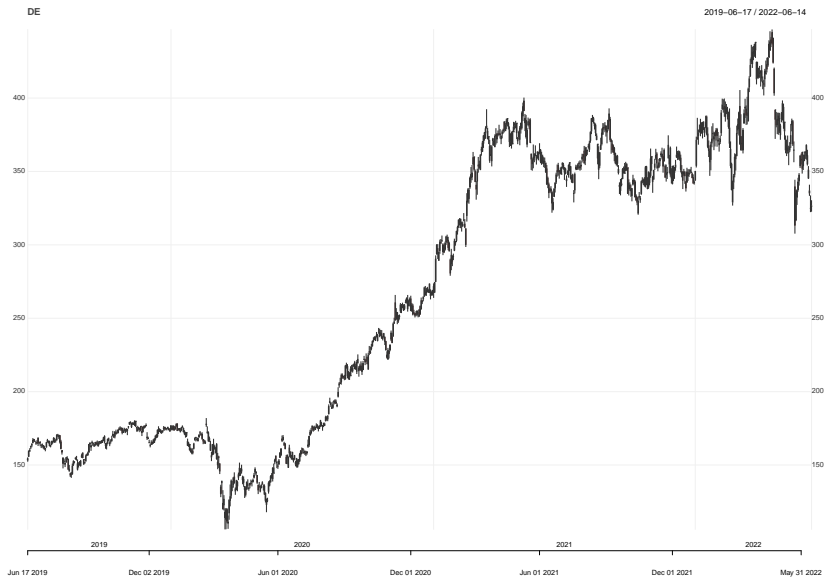
Pick a Stock | Example: John Deere

```
# Load {tidyquant}
library(tidyquant)

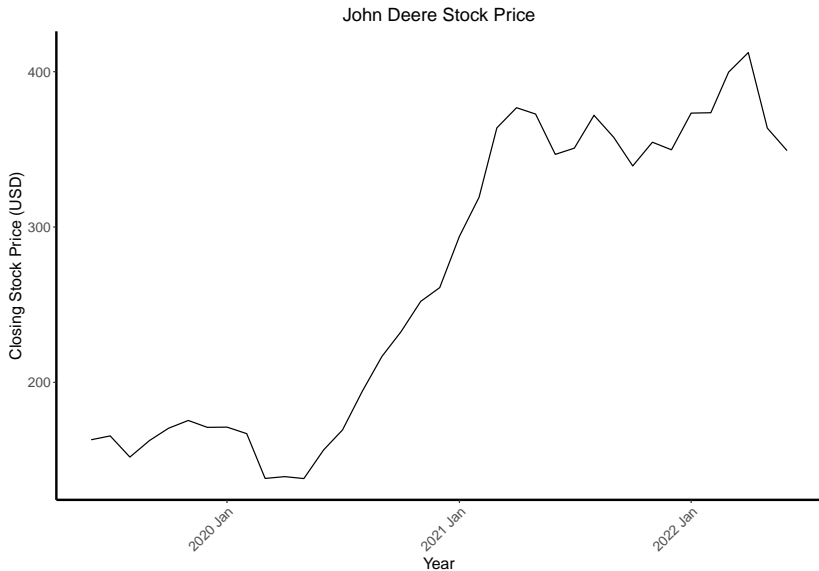
# Get data
getSymbols(
  "DE", # ticker symbol
  from = "2019-06-15", # start date
  to = "2022-06-15", # end date
  auto.assign = TRUE # assign variable 'XOM'
)

# First six rows of data
head(DE)
```

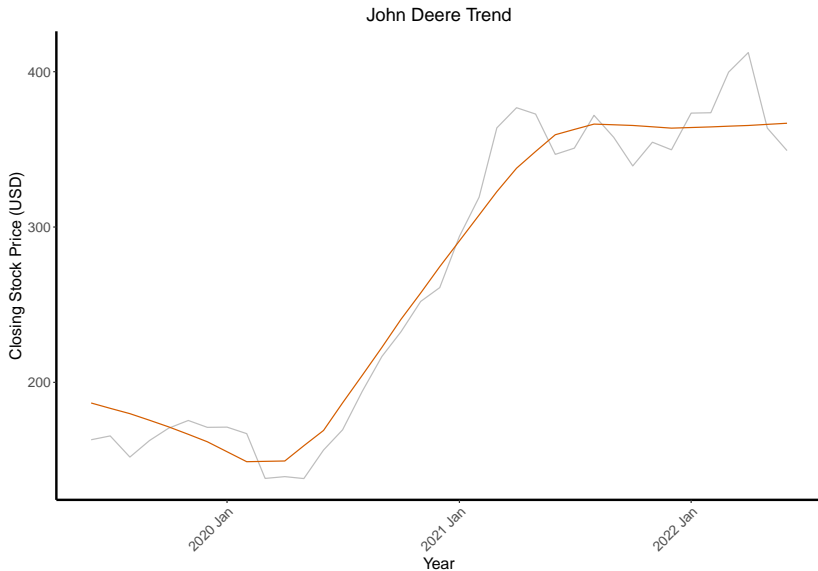
Stock Example | John Deere Chart



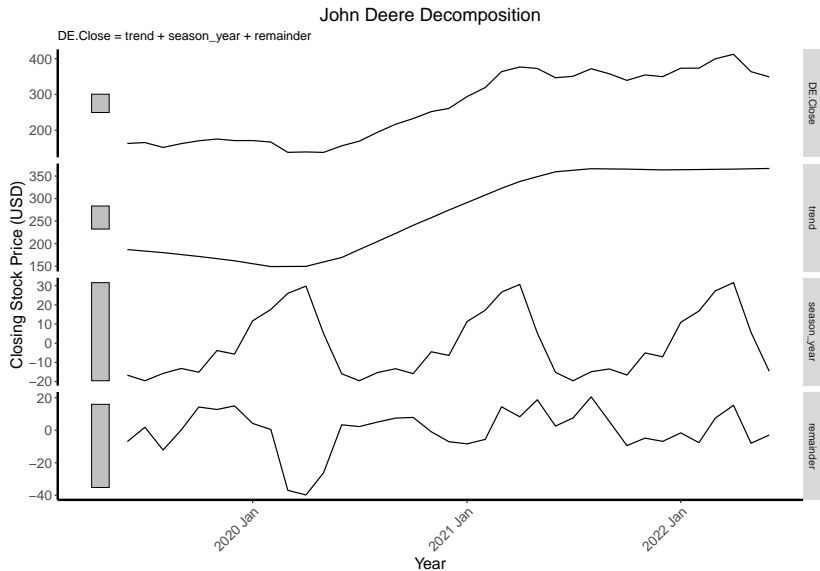
Stock Example | John Deere Closing Price



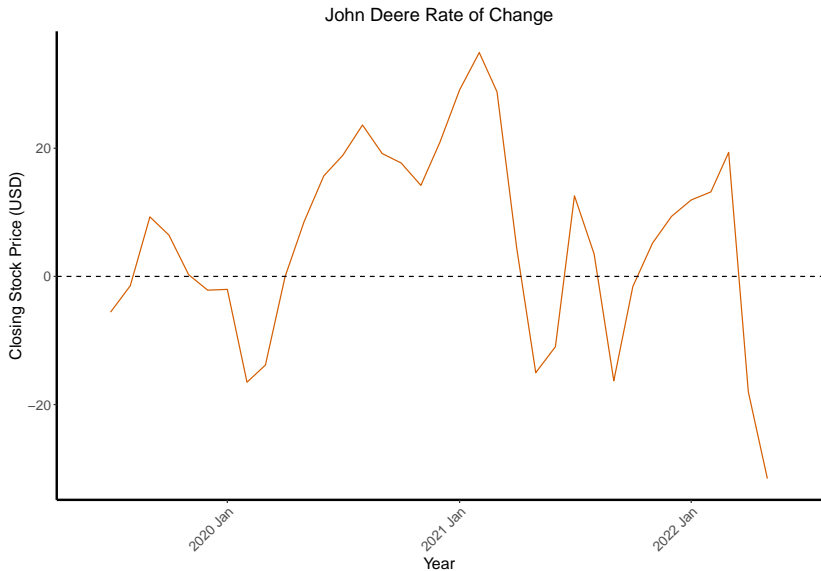
Stock Example | John Deere Trend



Stock Example | John Deere Decomposition



Stock Example | John Deere Changing Over Time



Stock Example | John Deere Cross-Validation

.model	RMSE	MAE	MASE	RMSSE	ACF1
Damped	20.864	16.788	0.167	0.168	0.069
Holt	22.861	19.013	0.189	0.184	0.146
SES	21.727	18.450	0.184	0.175	0.277

Stock Example | John Deere Model Fit

Series: DE.Close

Model: ETS(A,Ad,N)

Smoothing parameters:

alpha = 0.9998998

beta = 0.3526345

phi = 0.8000007

Initial states:

l[0] b[0]

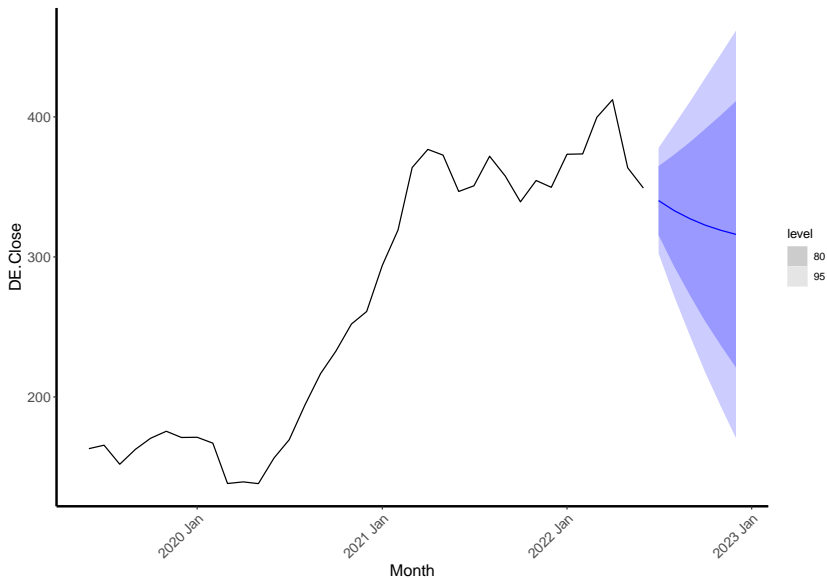
166.426 3.868474

sigma²: 369.625

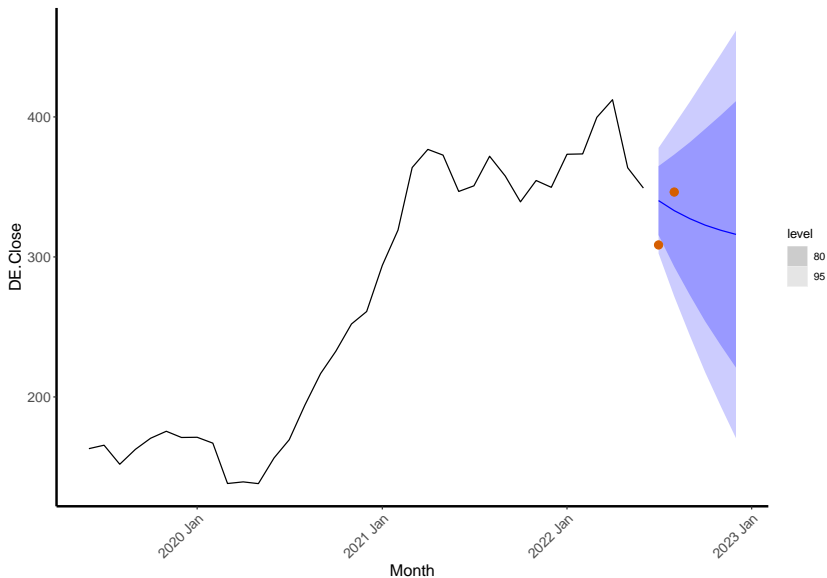
AIC	AICc	BIC
-----	------	-----

358.9943	361.7943	368.6598
----------	----------	----------

Stock Example | John Deere Prediction



Stock Example | John Deere Accurate?



For next week

- Pick any stock you're interested in (except for Google)
- Email me your stock name and ticker (e.g., John Deere; DE):
alexander.christensen@vanderbilt.edu
- Due next Monday (08.29.2022)