Scouting New Trails: An Algorithmic Approach to Teaching and Learning Statistical Modeling

Nicola Justice
Pacific Lutheran University

In collaboration with Andrew Zieffler, Robert delMas & Michael D. Huberty,
University of Minnesota

Collaborators

MacDonald, Peter. (2010). Andrew Zieffler (Minnesota). Licensed under Creative Commons 2.0. (Link to the license.) Retrieved from https://www.flickr.com/photos/ssc_liaison/42824717481

Andrew Zieffler

Bob delMas

Mike Huberty

The research I want to share with you today is…

- Exploratory

The research I want to share with you today is…

- Exploratory
- The best research I've ever done
    - Innovative
    - Potential to make positive change
    - Challenging for the statistics education community
        - Changes in curriculum?
        - Changes in power?

The research I want to share with you today is…

- Exploratory
- The best research I've ever done
    - Innovative
    - Potential to make positive change
    - Challenging for the statistics education community
        - Changes in curriculum?
        - Changes in power?
- Has had the most "push back" from the research community
    - Rejected from journals
    - Reviewer comments indicate misunderstanding, not proper critique

GOALS for this talk

- Introduce Algorithmic Modeling
    - Motivation (why?!)
    - Compare to Probabilistic Modeling (briefly)
    - Example using CART algorithm

GOALS for this talk

- Introduce Algorithmic Modeling
    - Motivation (why?!)
    - Compare to Probabilistic Modeling (briefly)
    - Example using CART algorithm
- Share our research introducing high school teachers to CART
    - What came easily
    - What was difficult

GOALS for this talk

- Introduce Algorithmic Modeling
  - Motivation (why?!)
  - Compare to Probabilistic Modeling (briefly)
  - Example using CART algorithm
- Share our research introducing high school teachers to CART
  - What came easily
  - What was difficult
- Discuss Implications & Questions

The context:

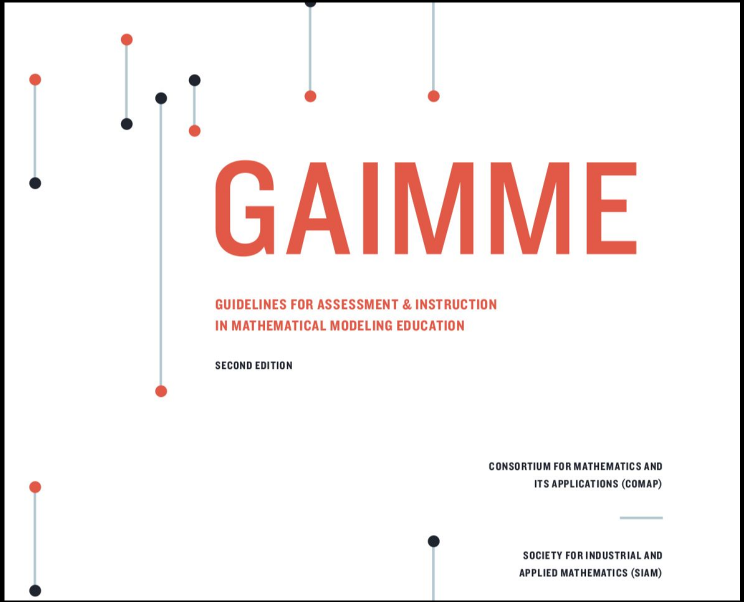
Statistical Modeling is….

- Important

GAIMME (2016)

https://www.siam.org/Portals/0/Publications/Reports/GAIMME_2ED/GAIMME-2nd-ed-final-online-viewing-color.pdf

- Modeling is important from Pre-K through college levels
- Principles of Modeling
  - Open-ended and messy
  - Students make genuine choices



Common Core State Standards (2010)

http://www.corestandards.org/Math/

- Decide if a specified model is consistent with results from a given data-generating process…
- Use data from a sample survey to estimate a population mean or proportion…
- Use data from a randomized experiment to compare two treatments… decide if differences between parameters are significant.

The Context:

Statistical Modeling is....

- Important
- Difficult for students (and teachers) to understand

The context:

Statistical Modeling is....

- Important
- Difficult for students (and teachers) to understand
- Difficult to define
    - Model building
    - Model evaluation

The Context:

Statistical Modeling is….

- Important
- Difficult for students (and teachers) to understand
- Difficult to define
    - Model building
    - Model evaluation
- Almost ubiquitously taught…
    - Using probabilistic models
    - In the context of statistical inference

More Context:

Computing power brings…

- Ability to analyze larger & more complex data sets
- Questions about the statistics curriculum
- New tools for statistical modeling
- Divide between academia & industry

Enter Brieman (2001)
Statistical Modeling: The Two Cultures

- *Probabilistic* modeling historically prevails in teaching / academia
- *Algorithmic* modeling prevails in industry

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony (avoid overfit)
- Goal: explanation—learning about the process

(Brieman, 2001)

Traditional STAT 101 Methods use Probabilistic Models

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

Traditional STAT 101 Methods use Probabilistic Models

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

$y_{ji} = \mu + \varepsilon_{ji}$

$y_{ji} = \mu + \alpha_j + \varepsilon_{ji}$

For $j \in \{1,2\}$,

$i \in \{1,2,\ldots n_j\}$

Traditional STAT 101 Methods use Probabilistic Models

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

$y_{ji} = \mu + \varepsilon_{ji}$

$y_{ji} = \mu + \alpha_j + \varepsilon_{ji}$

- Components: structural & random
- Tension: complexity & parsimony (avoid overfit)
- Goal: explanation—learning about the process

For $j \in \{1,2\}$,
$i \in \{1,2,\ldots n_j\}$

(Brieman, 2001)

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony (avoid overfit)
- Goal: prediction

Our Focus:
Classification (Decision) Tree Algorithmic Models

- Binary Response Variable
  - e.g., in the tragic sinking of the RMS *Titanic* (1912), each passenger died or survived.
- Observations are "classified" into one of the two possible outcomes
  - e.g., the model *predicts* whether a passenger died or survived
- The classification may or may not be correct

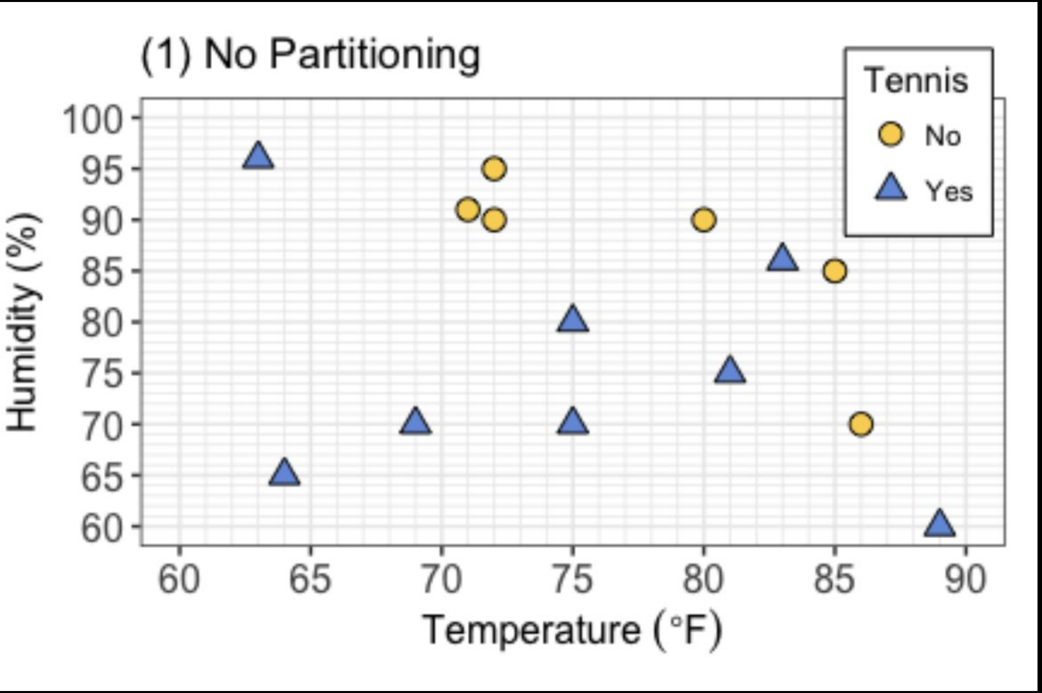
Example of Building an Algorithmic Model (modified from Witten, Frank and Hall 2011) using CART Algorithm by Breiman et al. (Strobl 2013)

- Goal: predict whether play tennis

- Two predictor variables

- If all cases classified as "Yes", accuracy = 57.1%

- Baseline or "no decision" model.



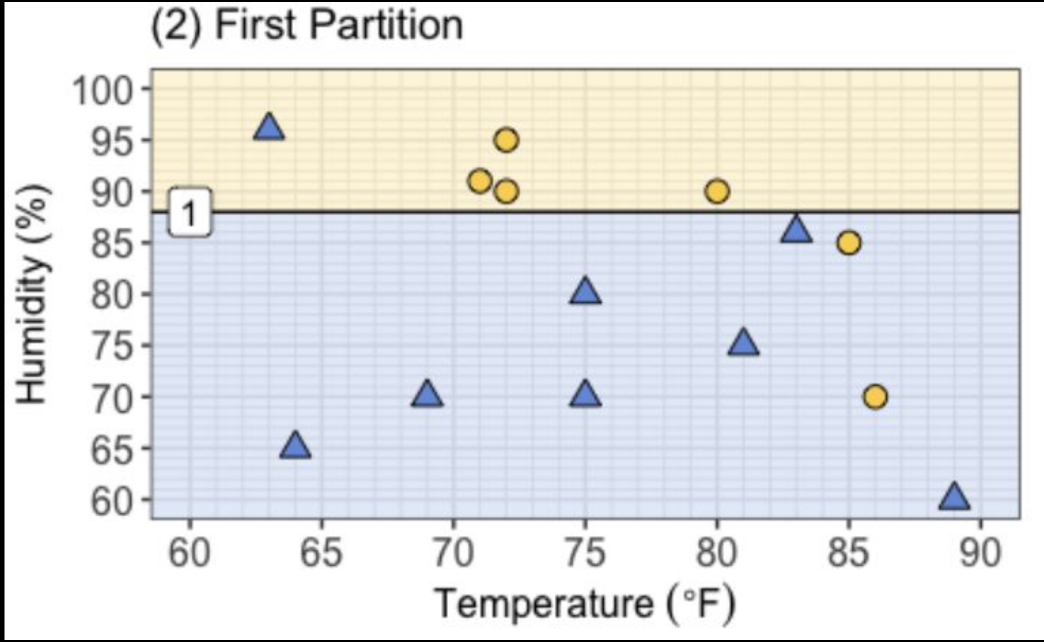
Adding one partition

- No single partition perfectly classifies all the cases

- Best scenario: partition between 85 & 90% humidity

- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6% accuracy

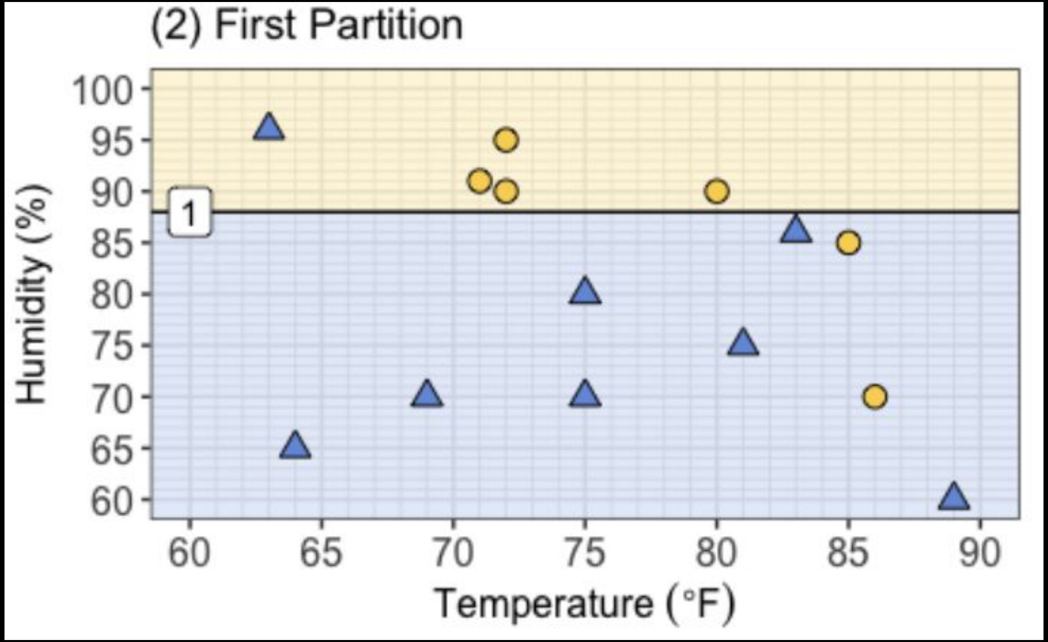
- “One partition” model has higher classification accuracy in exchange for greater complexity



Adding one partition

```
IF (Humidity < 88%)
THEN {Predict: Play tennis}
ELSE {Predict: Do not play tennis}
```

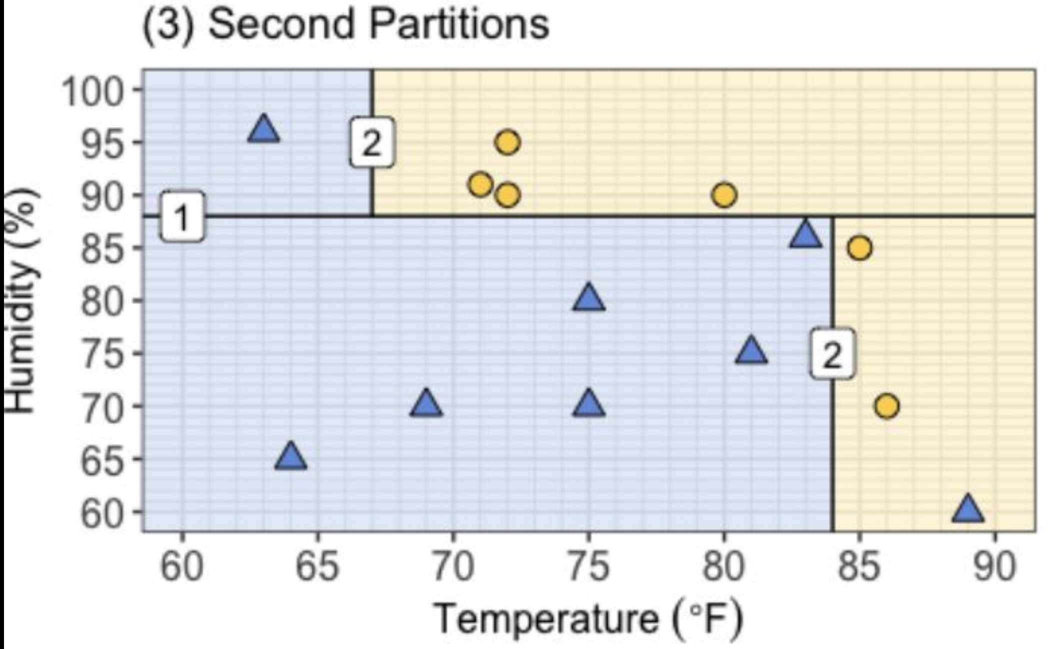
- No single partition perfectly classifies all the cases
- Best scenario: partition between 85 & 90% humidity
- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6% accuracy
- “One partition” model has higher classification accuracy in exchange for greater complexity



Second Level Partitions

- Partition again to optimize within each of the 1st-level partitions

- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6%
  - 2nd level partitions: 92.9%



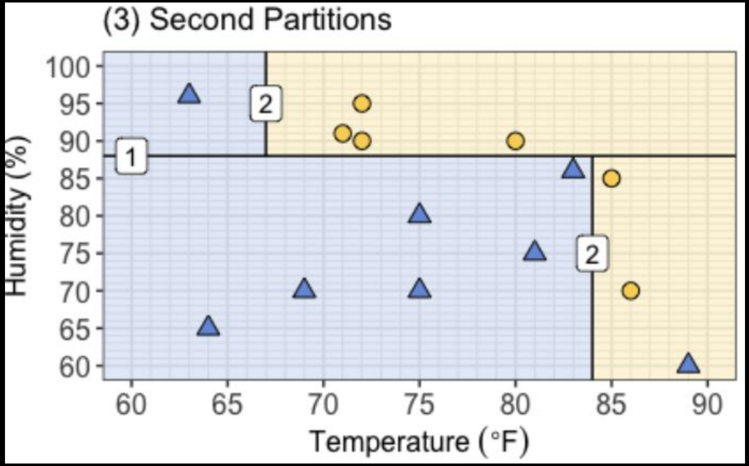
Second Level Partitions

- Partition again to optimize within each of the 1st-level partitions

- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6%
  - 2nd level partitions: 92.9%

- Higher classification accuracy in exchange for greater complexity

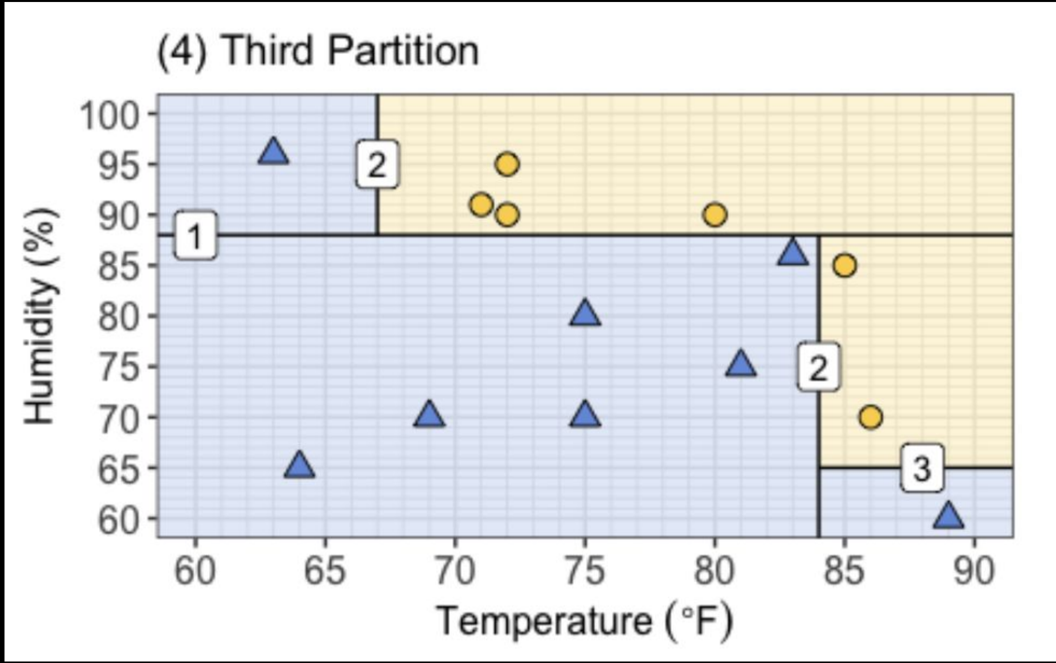
```
IF (Humidity < 88%)
THEN {
IF (Temperature < 84 degrees)
     THEN {Predict: Play tennis}
     ELSE {Predict: Do not play tennis}
}
ELSE {
IF (Temperature < 67 degrees)
THEN {Predict: Play tennis}
ELSE {Predict: Do not play tennis}
}
```



Third Level Partition

- Partition again to correctly classify all cases

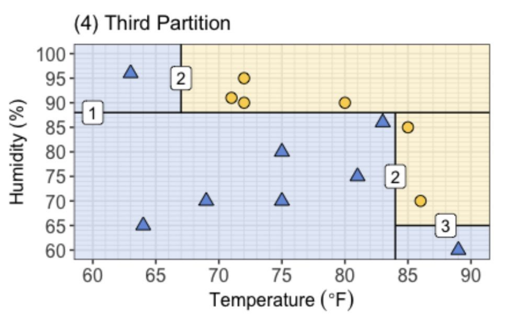
- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6%
  - 2nd level partitions: 92.9%
  - 3rd level partitions: 100%



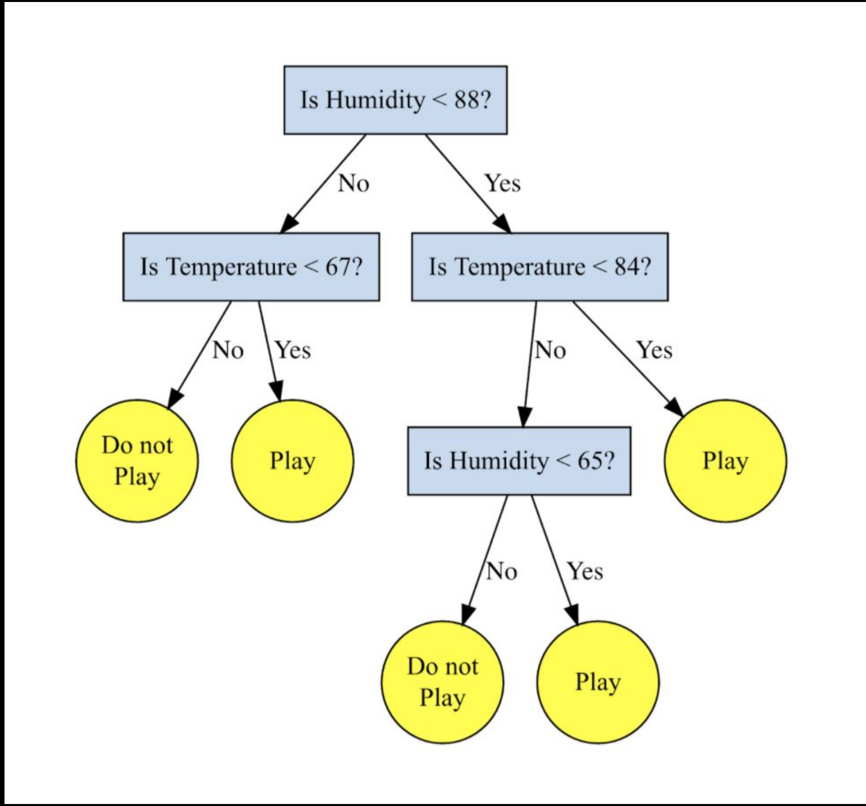
Third Level Partition

- Partition again to correctly classify all cases
- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6%
  - 2nd level partitions: 92.9%
  - 3rd level partitions: 100%
- Higher classification accuracy in exchange for greater complexity

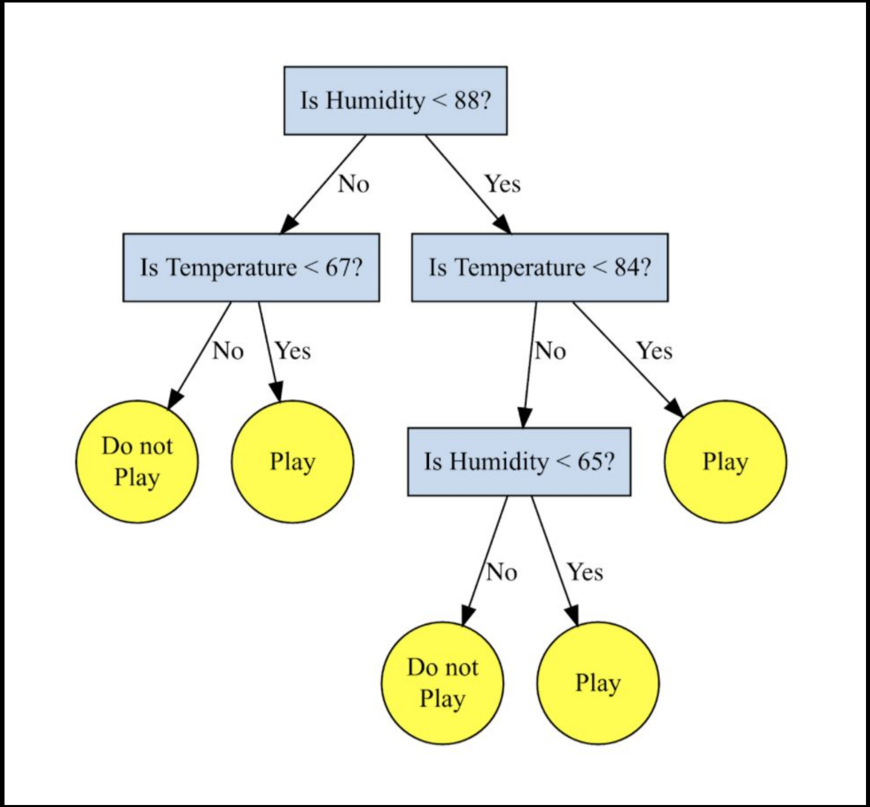
```
IF (Humidity < 88%
THEN {
IF (Temperature < 84 degrees)
     THEN {Predict: Play tennis}
ELSE {
IF (Humidity < 65%)
THEN {Predict: Play tennis}
ELSE {Predict: Do not play tennis}
}
}
ELSE {
IF (Temperature < 67 degrees)
THEN {Predict: Play tennis}
ELSE {Predict: Do not play tennis}
}
```



Depicting the Model using a Decision Tree



Depicting the Model using a Decision Tree



Suppose on a new occasion the friends do NOT play tennis when

- 66 °F
- 90% humidity

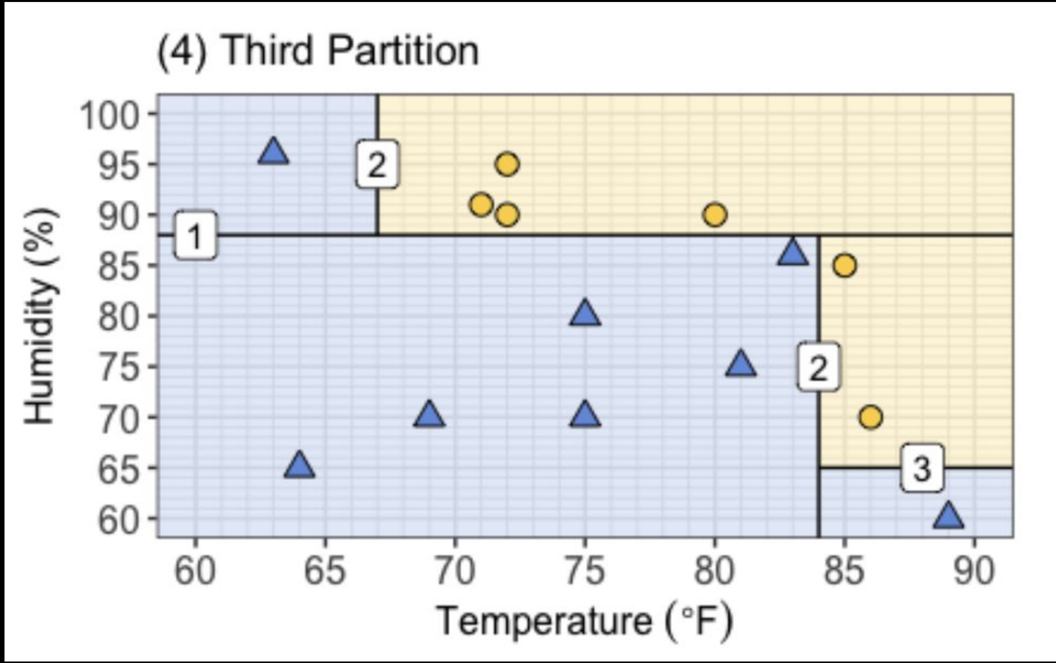
What does the model predict?

Is the model correct?

Model Evaluation

Model Evaluation

- Is the final model the best choice?
- Goal is prediction
- Would the model be just as accurate if some partitions were omitted?



Model Evaluation

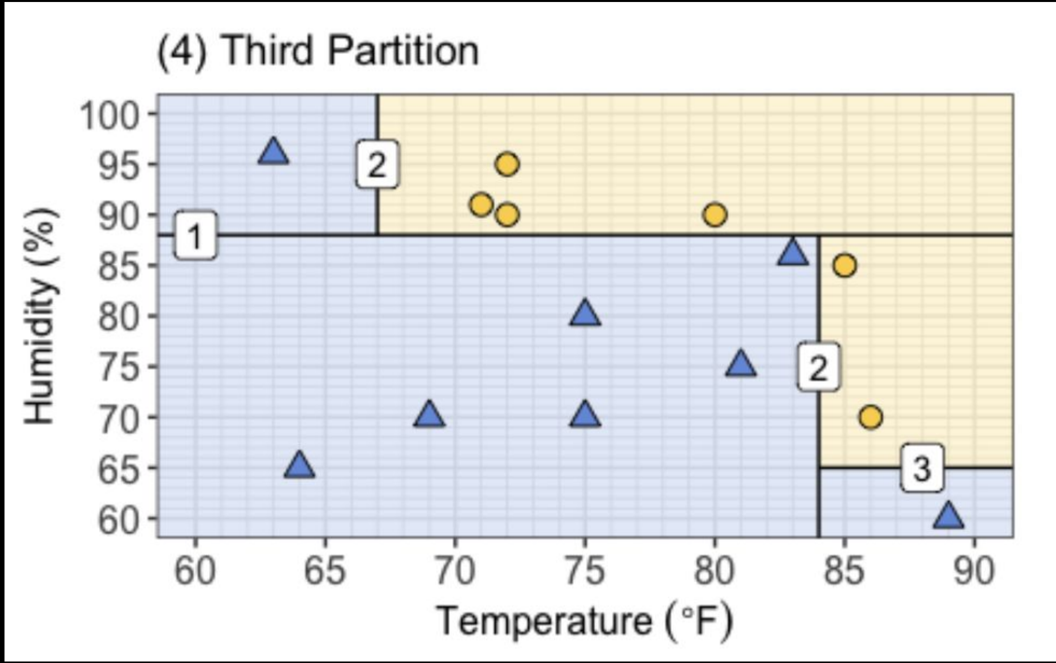
Tree Pruning:

- Use a hold-out set of data

  OR

- Use *a priori* thresholds of accuracy improvement

- Remove "branches" of the model if criteria not met



Tree Pruning

- Suppose we use a 10% improvement threshold

- Comparing accuracy:
    - Baseline: 57.1 %
    - One partition: 78.6%
    - 2nd level partitions: 92.9%
    - 3rd level partitions: 100%

3rd level provides only 7.1% improvement in accuracy…

…So we prune the 3rd level rule.

Tree Pruning

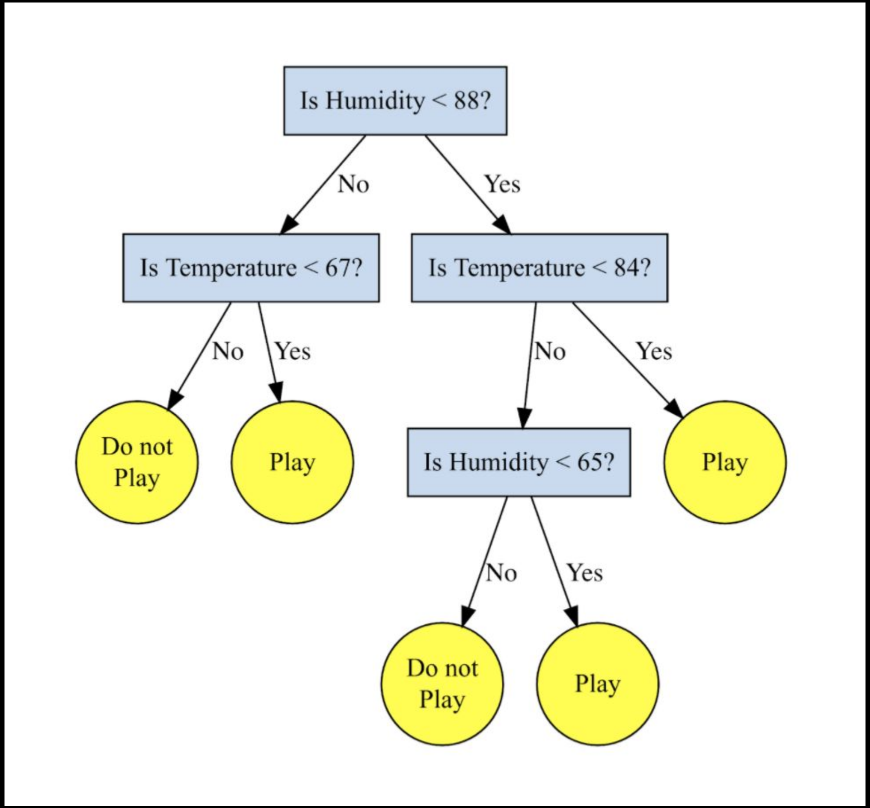
- Suppose we use a 10% improvement threshold

- Comparing accuracy:
  - Baseline: 57.1 %
  - One partition: 78.6%
  - 2nd level partitions: 92.9%
  - ~~3rd level partitions: 100%~~

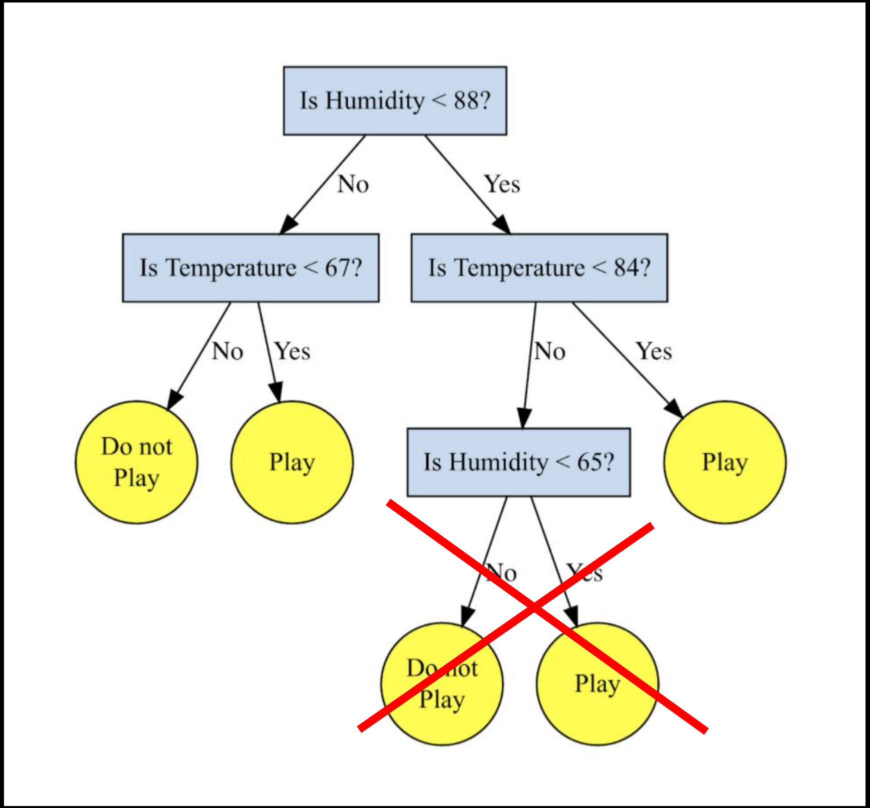
….Continue working from the bottom up until the 10% threshold is reached…

then stop pruning that "branch."

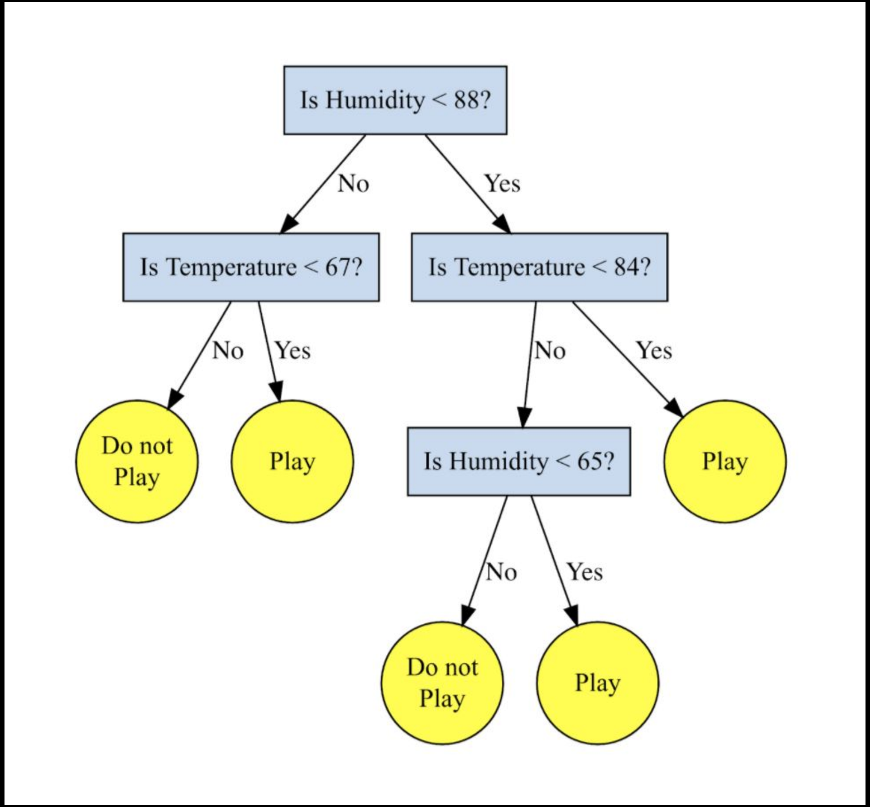
Unpruned model



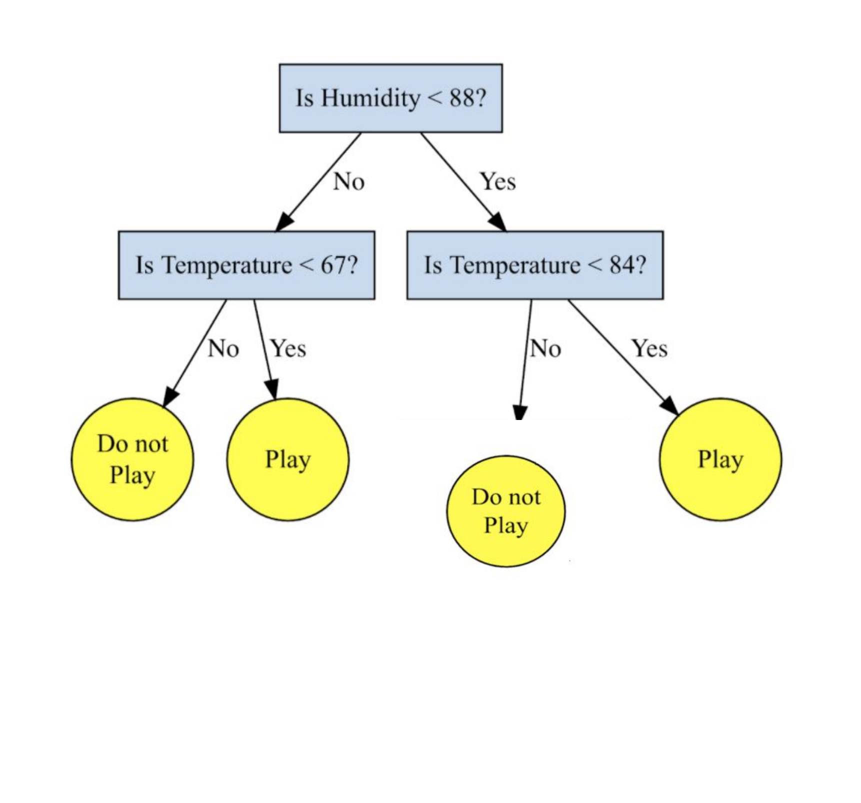
Pruned model



Unpruned model



Pruned model



Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

BONUS:

- Use computational thinking
  - important educational outcome (Wing, 2006)
  - serves a diverse audience (Weintrop et al. 2016)

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

BOTH ARE IMPORTANT!

BONUS:

- Use computational thinking
  - important educational outcome (Wing, 2006)
  - serves a diverse audience (Weintrop et al. 2016)

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

Taught in K-12 and intro college stats

BONUS:

- Use computational thinking
  - important educational outcome (Wing, 2006)
  - serves a diverse audience (Weintrop et al. 2016)

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

BONUS:

- Use computational thinking
  - important educational outcome (Wing, 2006)
  - serves a diverse audience (Weintrop et al. 2016)

Almost never taught in K-12 or intro stats

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

BOTH ARE IMPORTANT!

BONUS:

- Use computational thinking
  - important educational outcome (Wing, 2006)
  - serves a diverse audience (Weintrop et al. 2016)

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

BONUS:

- Use computational thinking
  - important educational outcome (Wing, 2006)
  - serves a diverse audience (Weintrop et al. 2016)

NOTICE similarities!

OK, so should algorithmic models be in the curriculum somewhere?

OK, so should algorithmic models be in the curriculum somewhere?

Would knowing about one type help with understanding the other?
....Or will it just confuse?

Research Question

To what extent can a curated set of professional development activities help secondary statistics teachers learn about algorithmic modeling?

Research Questions, more specifically

1. To what extent do the professional development activities help secondary statistics teachers **understand important ideas related to algorithmic modeling,** especially those related to overfit and the need for cross-validation to prevent overfitting?

2. Do concepts and methods used for **probabilistic modeling interfere** with secondary statistics teachers' understanding of concepts and methods used for algorithmic modeling? If so, what is the nature of the interference?

College in the Schools (CIS) Program

- Students take university course in high school
- Taught by high school mathematics teachers
- Part of our job is to keep the CIS teachers current

Participants: CIS Teachers (n = 11)

- Minimal prior training in statistics
- Bachelor's or Master's degree in Mathematics or Mathematics Education
- Some previously taught Advanced Placement (AP) Statistics
- Teaching CIS Statistics for 1–3 years

The Professional Development

- Designed lessons to introduce algorithmic modeling
    - Lessons: small groups of 2-4 teachers
    - Discussions: large group of 11 teachers + facilitators
    - Reflections: individually

The Professional Development

5 days of professional development & assignments

3 Days (End of Summer)

ACTIVITIES
- Email SPAM Classification
- Titanic Classification Trees I
- Titanic Classification Trees II

ASSIGNMENT
- Individual Reflection

1 Day (Fall)

ACTIVITIES
- Building Decision Trees
- Building and Evaluating an ‘Optimal’ Decision Tree

ASSIGNMENT
- Individual Follow-Up

1 Day (Spring)

ACTIVITY
- Recursive Partitioning

Data

- Video Recordings of the teachers working on the activities
- Video Recordings of the large-group discussions after the activities
- Copies of all the teachers' work on the activities
- Responses to reflection questions

Analysis

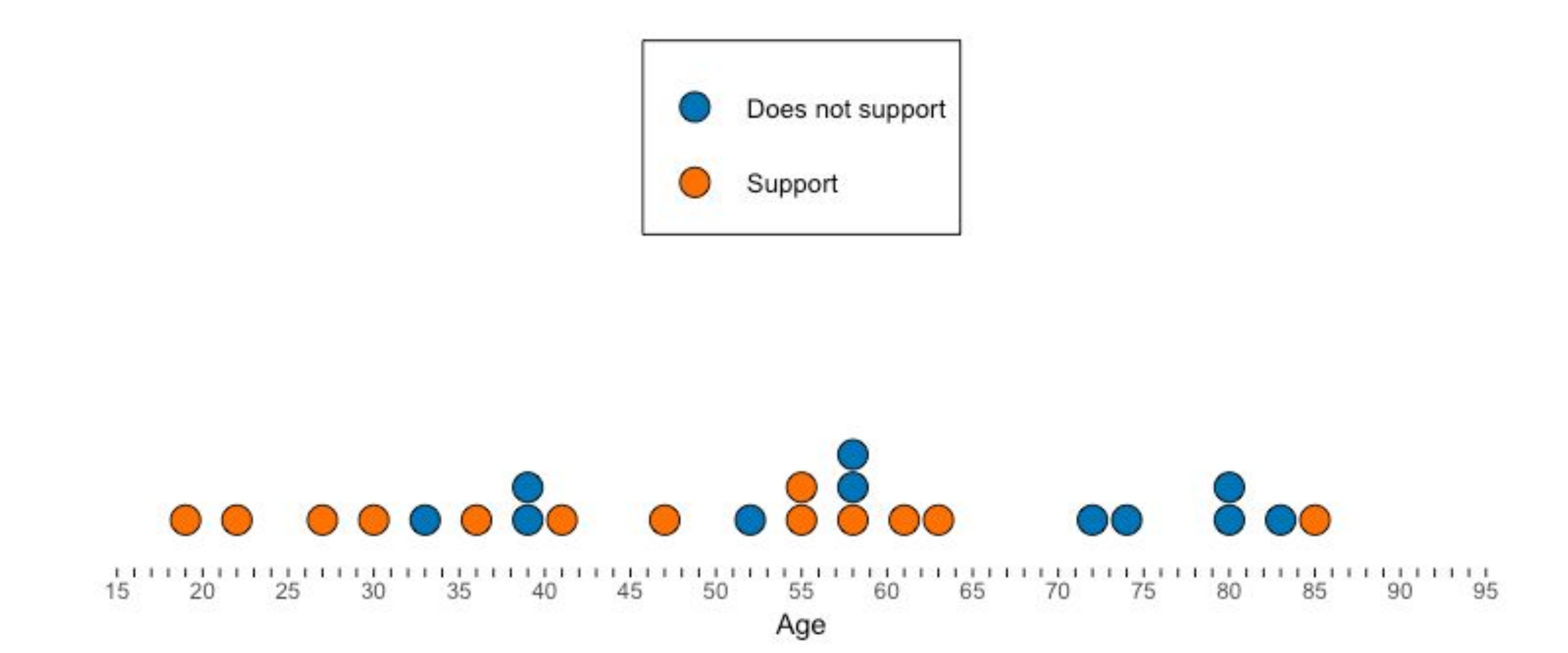
- Grounded Theory Qualitative Study
- Viewed the data independently
- Met frequently to discuss what we saw
- Came to consensus on interpretations
- Looked for themes that emerged
- We hold the results loosely; this is exploratory!

Results

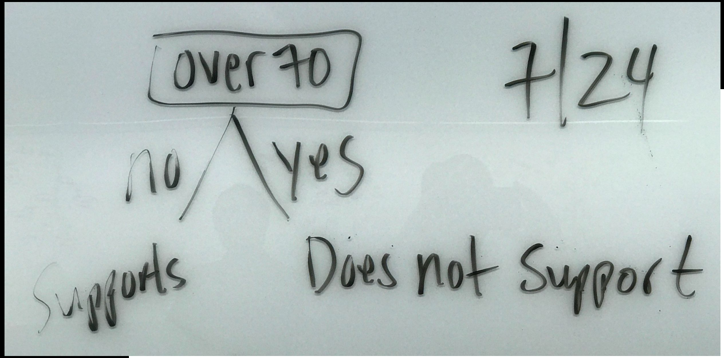
Teachers showed evidence of understanding many aspects of decision trees.

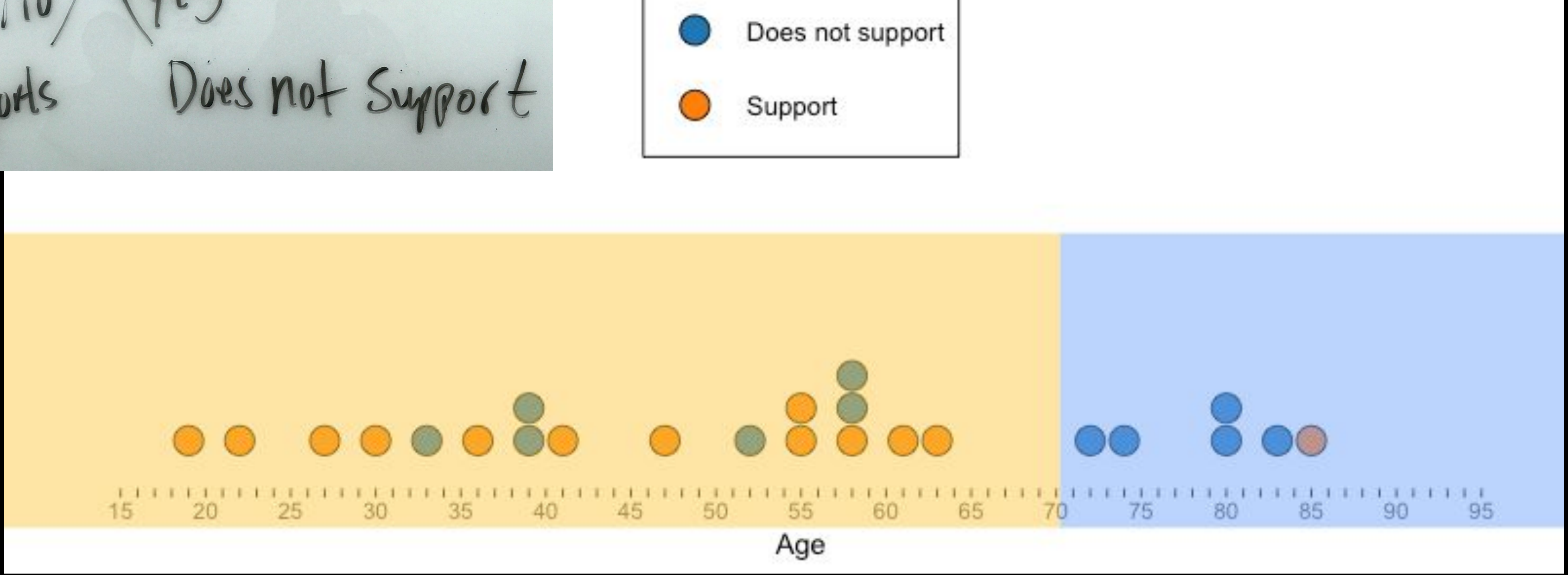
- Able to read a decision tree
- Could use the decision tree to classify cases
- Able to build a decision tree
- Sensitivity to overfit

Fitting a Decision Tree to Data

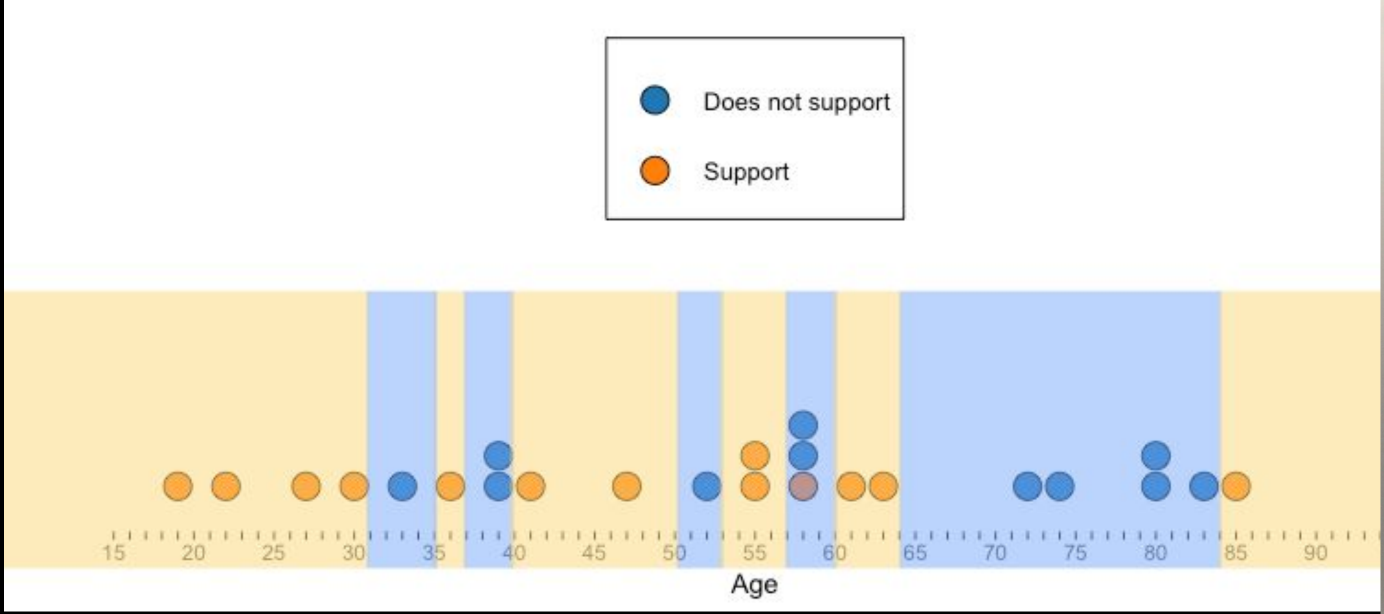


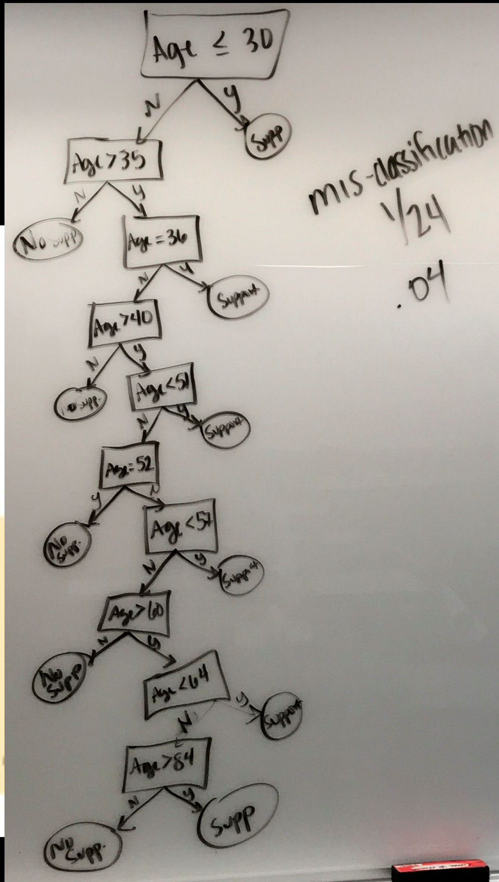
Fitting a Decision Tree to Data: A Simple Model





Fitting a Decision Tree to Data: Overfitted Model





Sensitivity to Overfit

Andy: Well, you got a darn good misclassification rate, right? And your model predicts really well.

Mark: For right now.

Katy: Except you're going to give us more data. And that's going to be.. we're going to wish we didn't do that [Laughter]….

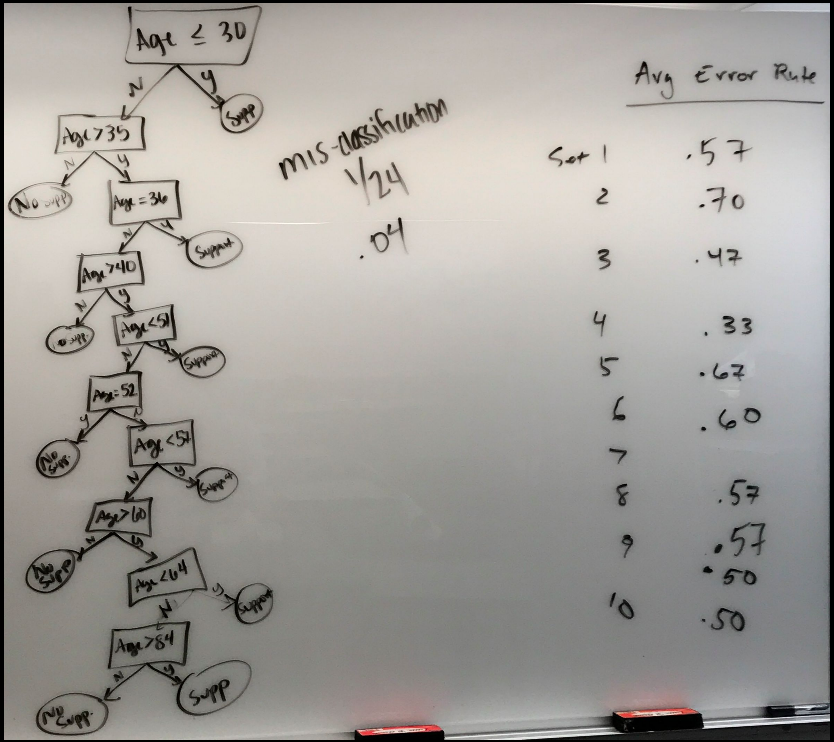
….

Barb: The rules are very specific to one data set. That doesn't mean they're good.

Promoting Sensitivity to Overfit: Cross Validation

The complex model had...

- low error rate for training set
- high error rates for 10 validation sets



Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

Possible Interference

- Perhaps conflation with *p*-values

Annie: Well, that would change it to 8% wrong. Which is still good. I think anything under 10% is good.

Barb: Yeah. Thinking about *p*-values.

Barb: Yeah, I think that, I just kind of heard him say it, kind of that what percent is acceptable? I mean, you kind of think of that *p*-value, that 5% is kind of that marker, acceptable, so anything less

Annie: So I think we're good. I mean, if you take it all together we're at 3%

More Possible Interference

- Used 50/50 model (not base rates) as their basis of comparison
- Used Absolute (not relative) standards of model performance

Comparison to 50/50 Model

Katy: …Was the tree algorithm a “good” model? …. 70%. So, it's a C-minus?

Mark: Yeah, we really don't have that standard, [?] it seems like it's [?]

Katy: It's better than half. Half and half.

More Possible Interference: “Tyranny of Context”

Kim: That's weird. If it's not, if it's a woman, we needed to know if they had a third class ticket?

Katy: Because they might not, they were on a lower level.

Kim: But why, why don't you ask, why did they not ask that for males?

Katy: Because most males died. They didn't get them on boats.

Kim: Oh, because of women and children first.

Mark: Who cares why?

More Possible Interference: "Tyranny of Context"

Kim: I want to understand the model, Mark!

Kim: So, then, if they had a first class ticket, they were below. They wanted to know older than 16. So, if they were young, they survived, because the children went. Okay, and they're older, more than three immediate family members aboard, if they were, so what, if you were a woman with a third class ticket, and you were older than 16 and you had a lot of kids with you, you died, because you didn't get on the boats? If you weren't, they want to know if you're older than 28, you died? And if you were, holy moly! And if you were, what was the last one? So, if you were between 22 and 28, you survived?

Katy: According to this model.

Kim: According to this model. Wow!

Katy: Maybe they're assuming that the strong years of your life, you're able to survive in the water a little better, if you were in the water.

Katy: That's crazy. Okay.

Algorithmic Models

- Components: set of rules
- Tension: complexity & parsimony
- Goal: prediction

Probabilistic Models

- Components: structural & random
- Tension: complexity & parsimony
- Goal: explanation

Summary

- In our study, the teachers seemed to pick up many aspects of algorithmic modeling fairly easily
  - Could read, interpret, and create decision trees
  - Showed some understanding of overfit

Summary

- In our study, the teachers seemed to pick up many aspects of algorithmic modeling fairly easily
  - Could read, interpret, and create decision trees
  - Showed some understanding of overfit
- Some notions of probabilistic modeling may have interfered
  - Absolute standards instead of relative (baseline) comparison
  - Tyranny of context

We Have New Research Questions

- Would it benefit students if algorithmic modeling were introduced earlier in the K-12 curriculum? … before probabilistic modeling?

- How might early introduction to algorithmic modeling conflate or support students' understanding of probabilistic modeling?

Traditional STAT 101 Methods use Probabilistic Models

Ho: $\mu_1 = \mu_2$

Ha: $\mu_1 \neq \mu_2$

$y_{ji} = \mu + \varepsilon_{ji}$

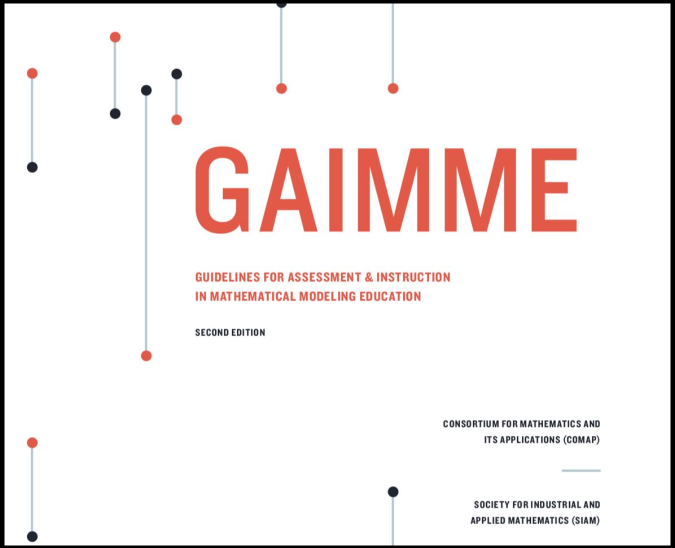
$y_{ji} = \mu + \alpha_j + \varepsilon_{ji}$

For $j \in \{1,2\}$,
$i \in \{1,2,\ldots n_j\}$

GAIMME (2016)

https://www.siam.org/Portals/0/Publications/Reports/GAIMME_2ED/GAIMME-2nd-ed-final-online-viewing-color.pdf

- Modeling is important from Pre-K through college levels
- Principles of Modeling
  - Open-ended and messy
  - Students make genuine choices



Future Research

- Where in the curriculum might algorithmic modeling be included?
    - Before/after/in lieu of probabilistic modeling?
- Are there other precursor experiences that students/teachers need with algorithms that help develop reasoning?
- What is the appropriate amount of “coding” to introduce in a teaching sequence about algorithmic modeling?

More Future Research and Teaching Implications

- Can we re-sequence/edit the activities to improve participant understanding?
  - Create a new activity that highlights multivariate reasoning early in the sequence?
  - Remove context in any of the earlier activities?
  - Should ideas of overfit be moved to after focusing on the algorithm, or left integrated throughout from the beginning?
- What other activities/technology should be included?
  - Using software (e.g., CART package) to classify a larger data set?
  - Inclusion of continuous outcomes (regression trees)?

Thanks!

Thanks! … and References

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science, 16* (3), 199–231.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. I. (1984), *Classification and Regression Trees*, Belmont, California: Wadsworth.

Doerr, H. M. (2007). What knowledge do teachers need for teaching mathematics through applications and modelling? In W. Blum, P. L. Galbraith, H. W. Henn & M. Niss (Eds.), *Modelling and applications in mathematics education.* New York: Springer.

Garfunkel, S., & Montgomery, M. (2016). Guidelines for assessment and instruction in mathematical modeling education. Bedford, MA: Consortium for Mathematics and Its Applications and Society for Industrial and Applied Mathematics. Retrieved from http://www.siam.org/Portals/0/Publications/Reports/gaimme-full_color_for_online _viewing.pdf?ver=2018-03-19-115454-057

National Governors Association Center for Best Practices and Council of Chief State School Officers (2010). Common Core State Standards. Retrieved from http://www.corestandards.org/Math/.

Strobl, Carolin. (2013), "Data Mining," in *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2: Statistical Analysis*, ed., T. A. Little, Oxford University Press, DOI: 10.1093/oxfordhb/9780199934898.013.0029.

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25* (1), 127–147.

Wing, J. M. (2006). Computational thinking. *Communications of the ACM , 49* (3), 3335.

… and an Invitation

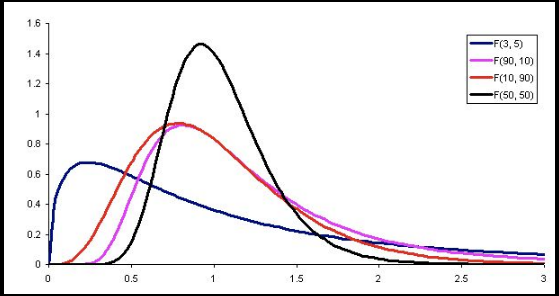
Contact: njustice@plu.edu

All the materials used in the professional development are available at: https://github.com/zief0002/srtl-11

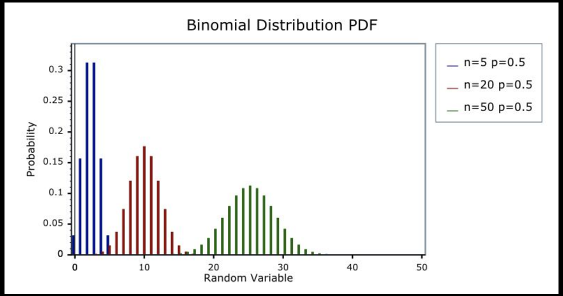
ADDITIONAL SLIDES FOR REFERENCE

Historically how Statistics and Probability *were* taught

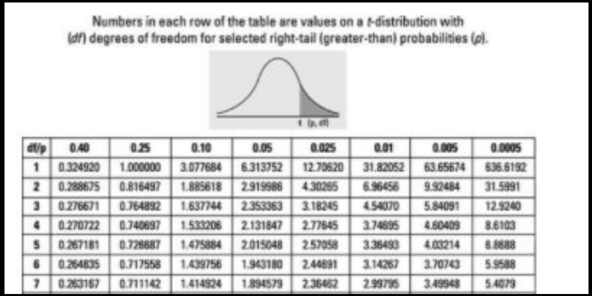
- A lot of consensus on “the introductory course”
    - Data Collection
    - Descriptive Statistics
    - Introductory Probability
    - Inferential Statistics using Probability Models



https://www.vosesoftware.com/riskwiki/Fdistribution.php



https://valelab4.ucsf.edu/svn/3rdpartypublic/boost/libs/math/doc/sf_and_dist/html/math_toolkit/dist/dist_ref/dists/binomial_dist.html



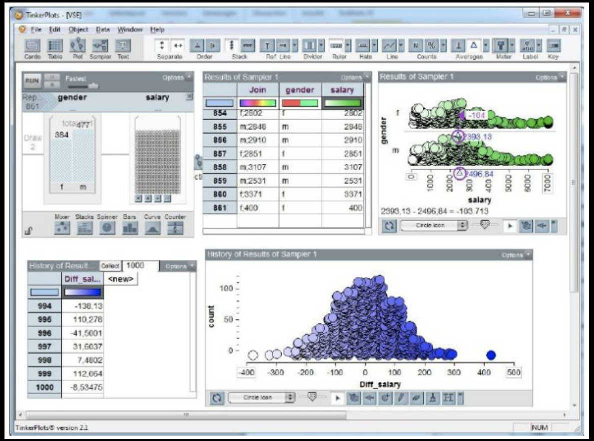
Numbers in each row of the table are values on a *t*-distribution with (*df*) degrees of freedom for selected right-tail (greater-than) probabilities (*p*).

df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079

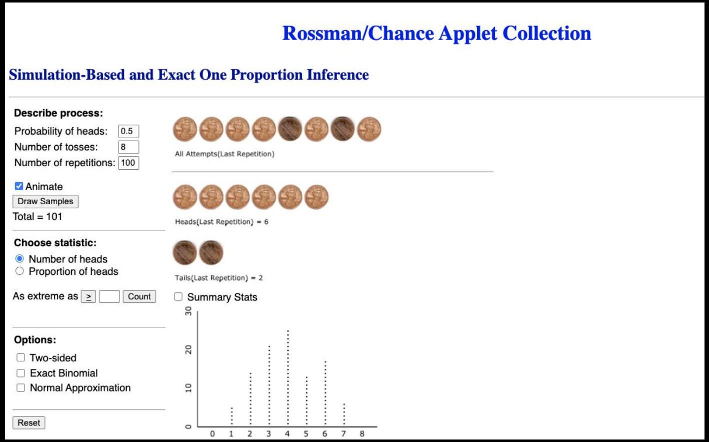
https://www.dummies.com/education/math/statistics/how-to-use-the-t-table-to-solve-statistics-problems/

More recently, there is less consensus

- Some instructors use “simulation based inference”
- Technology tools are used to generate probability models.



Biehler, R., Frischemeier, D., & Podworny, S. (2015). Preservice teachers reasoning about uncertainty in the context of randomization tests. *Reasoning about uncertainty:* In Zieffler, Z & Fry, E. (Eds). *Learning and teaching informal inferential reasoning*, 129-162.



http://www.rossmanchance.com/applets/OneProp/OneProp.htm

AlexiusHoratius. (2010). West side (rear) of the the Education Sciences Building at the Minneapolis campus of the University of Minnesota in the USA. Licensed under Creative Commons 3.0. (Link to the license.) Retrieved from https://commons.wikimedia.org/wiki/File:Education_Sciences_Building_Minnesota_6.jpg

University of Minnesota

- Ph.D. in Quantitative Methods in Education (Statistics Applied to Education Data)

How Students Learn Probability and Statistics

- It is VERY difficult for students to have authentic learning
  - Much like Physics
  - Might learn procedures on paper
  - It is very difficult for the realities to “set in”
  - E.g., gambler’s fallacy

Pixabay (2016) “Coins on Brown Wood.” Licenced under Pexels: free to use.
https://www.pexels.com/photo/antique-bills-business-cash-210600/

Research suggests:

- Students have trouble learning much of statistics
- Teachers & future teachers have trouble with probability & statistics!!

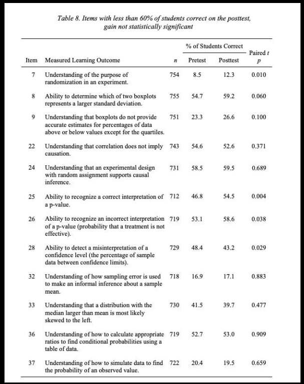
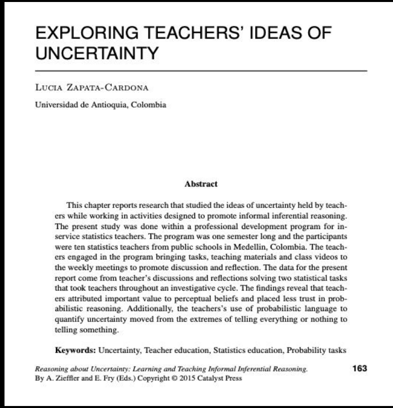


Table 8. Items with less than 60% of students correct on the posttest, gain not statistically significant

Item	Measured Learning Outcome	*n*	% of Students Correct: Pretest	% of Students Correct: Posttest	Paired *t* *p*
7	Understanding of the purpose of randomization in an experiment.	754	8.5	12.3	0.010
8	Ability to determine which of two boxplots represents a larger standard deviation.	755	54.7	59.2	0.060
9	Understanding that boxplots do not provide accurate estimates for percentages of data above or below values except for the quartiles.	751	23.3	26.6	0.100
22	Understanding that correlation does not imply causation.	743	54.6	52.6	0.371
24	Understanding that an experimental design with random assignment supports causal inference.	731	58.5	59.5	0.689
25	Ability to recognize a correct interpretation of a p-value.	712	46.8	54.5	0.004
26	Ability to recognize an incorrect interpretation of a p-value (probability that a treatment is not effective).	719	53.1	58.6	0.038
28	Ability to detect a misinterpretation of a confidence level (the percentage of sample data between confidence limits).	729	48.4	43.2	0.029
32	Understanding of how sampling error is used to make an informal inference about a sample mean.	718	16.9	17.1	0.883
33	Understanding that a distribution with the median larger than mean is most likely skewed to the left.	730	41.5	39.7	0.477
36	Understanding of how to calculate appropriate ratios to find conditional probabilities using a table of data.	719	52.7	53.0	0.909
37	Understanding of how to simulate data to find the probability of an observed value.	722	20.4	19.5	0.659

delMas, R., Garfield, J. B., Ooms, A., & Chance, B. (2007). Assessing Students' Conceptual Understanding After A First Course In Statistics. *Statistics Education Research Journal*, *6*(2).



EXPLORING TEACHERS' IDEAS OF UNCERTAINTY

LUCIA ZAPATA-CARDONA

Universidad de Antioquia, Colombia

Abstract

This chapter reports research that studied the ideas of uncertainty held by teachers while working in activities designed to promote informal inferential reasoning. The present study was done within a professional development program for in-service statistics teachers. The program was one semester long and the participants were ten statistics teachers from public schools in Medellin, Colombia. The teachers engaged in the program bringing tasks, teaching materials and class videos to the weekly meetings to promote discussion and reflection. The data for the present report come from teacher's discussions and reflections solving two statistical tasks that took teachers throughout an investigative cycle. The findings reveal that teachers attributed important value to perceptual beliefs and placed less trust in probabilistic reasoning. Additionally, the teachers's use of probabilistic language to quantify uncertainty moved from the extremes of telling everything or nothing to telling something.

Keywords: Uncertainty, Teacher education, Statistics education, Probability tasks

Reasoning about Uncertainty: Learning and Teaching Informal Inferential Reasoning. 163
By A. Zieffler and E. Fry (Eds.) Copyright © 2015 Catalyst Press

Zapata-Cardona, L. (2015), "Exploring Teachers' Ideas of Uncertainty," in Reasoning about Uncertainty: Learning and Teaching Informal Inferential Reasoning, eds., A. Zieffler and E. Fry, Minneapolis, MN: Catalyst Press, pp. 95–127.

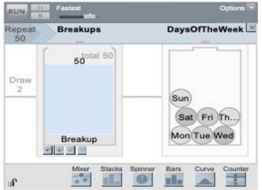


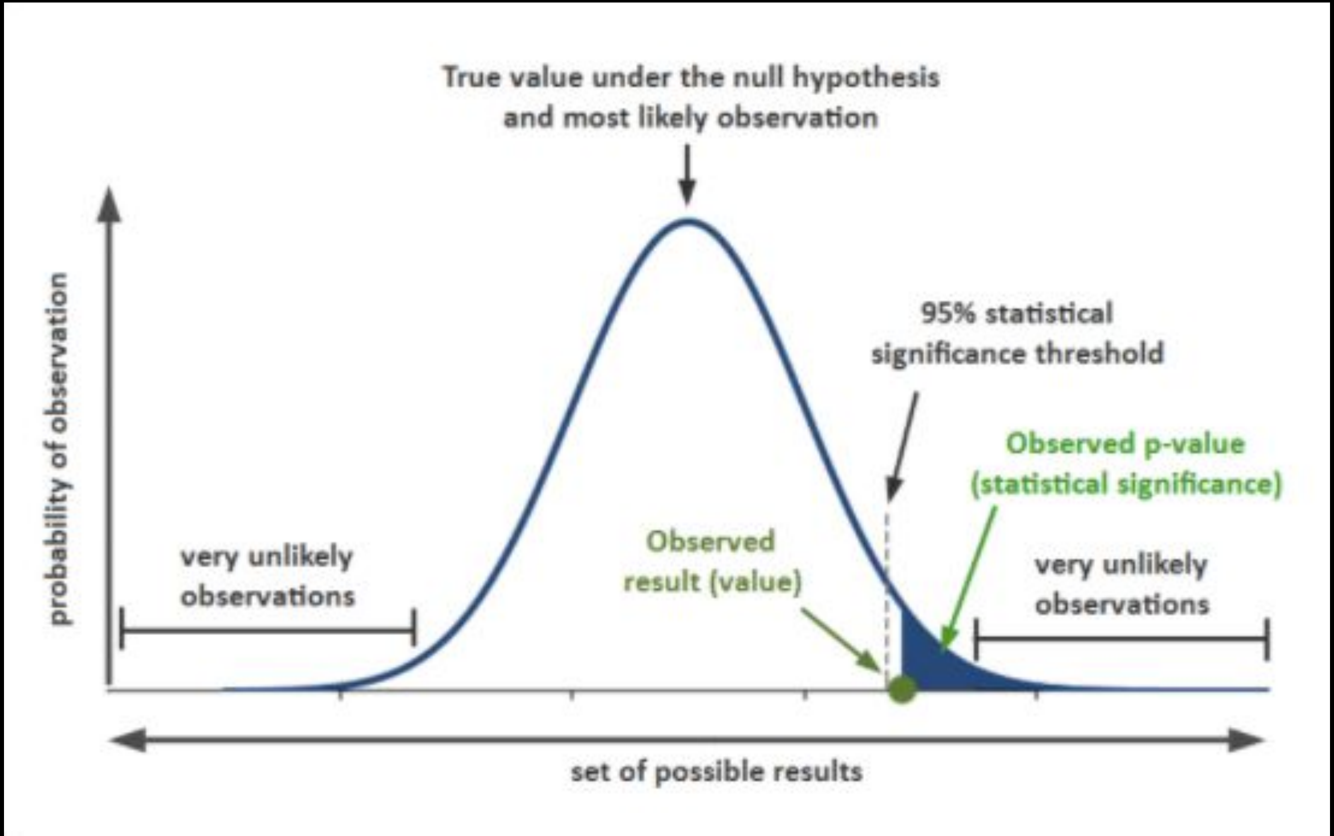
Figure 17. Student J's TinkerPlots™ model for the Facebook task

Student J: I set up a sampler with a stacked and mixer sampler. My stacked sampler was set to without replacement and labeled breakups with 50 stacks, and the mixer was set up with days of the week with replacement. I then set the sampler to repeat 50 times.

Figure 18. Student J's written explanation for model of the Facebook task

Noll, J., & Kirin, D. (2016). Student Approaches to Constructing Statistical Models using TinkerPlots TM. Technology Innovations in Statistics Education, 9(1).

Backwards thinking:



We can’t give up.

Modeling is Important!

Summary of the Problem:

- Modeling is important

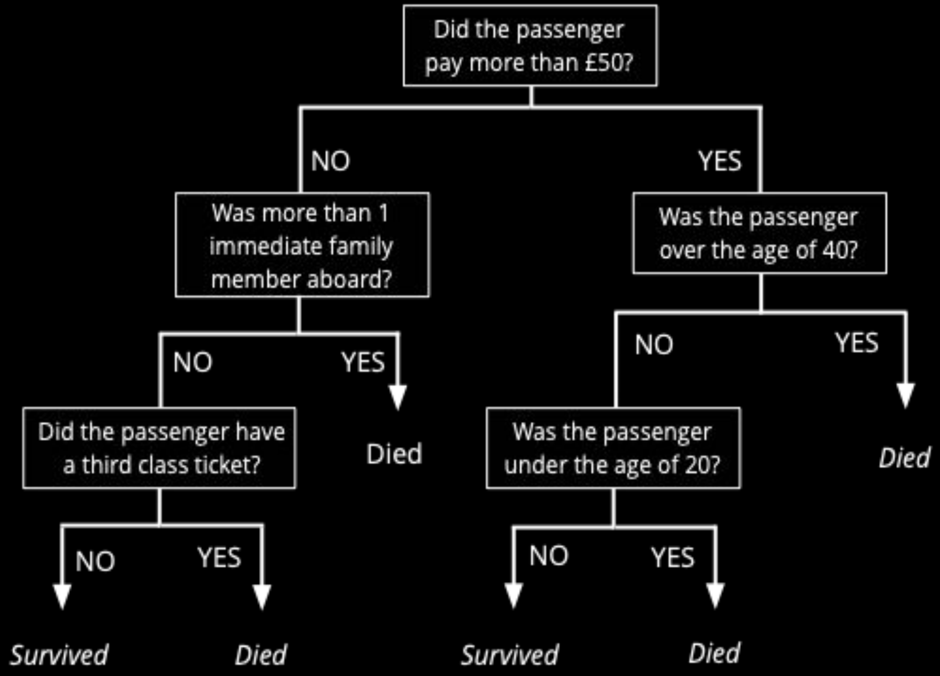
- With the current methods and context, evidence suggests it is extremely difficult to learn statistical modeling.

Is it Time to Put the CART Before the Horse? A Case for Algorithmic Modeling in the K–12 Curriculum

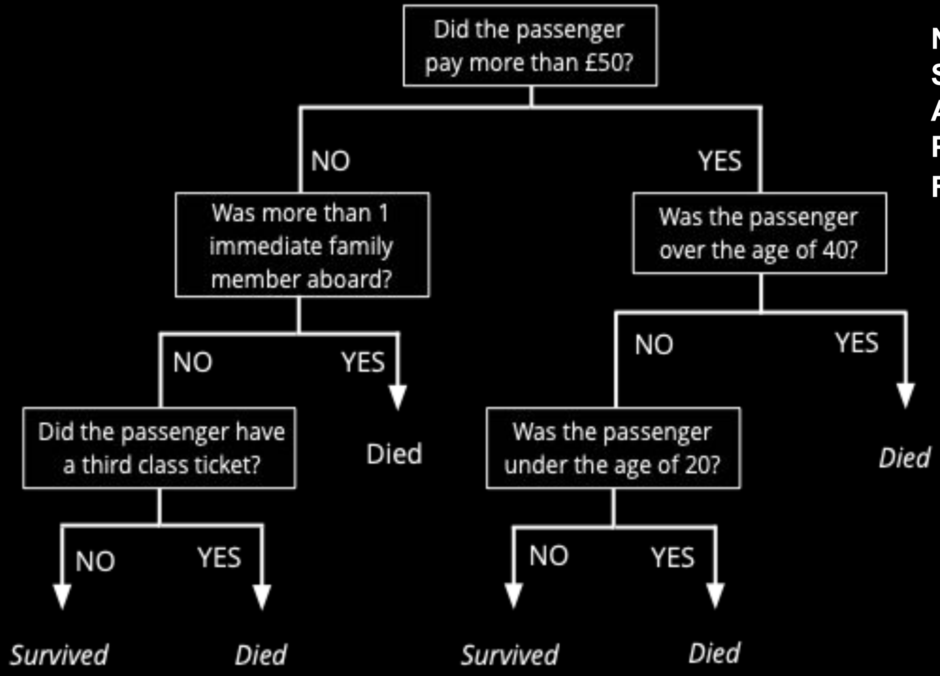
Nicola Justice
Pacific Lutheran University

Michael D. Huberty, Andrew Zieffler, & Robert delMas
University of Minnesota

Using a Decision Tree to Classify Cases



Using a Decision Tree to Classify Cases

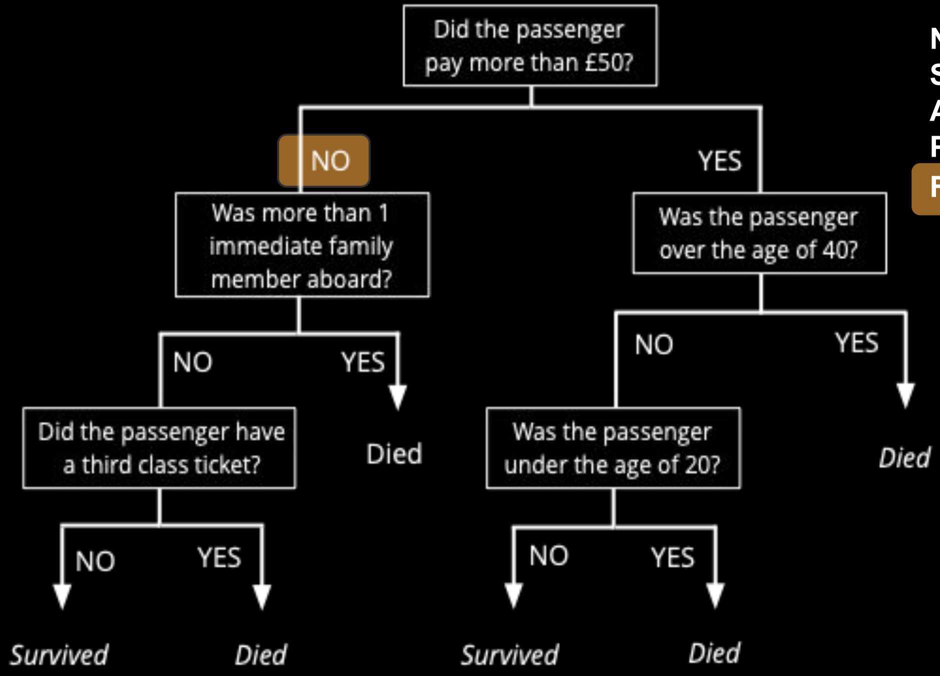


Name: Levy, Mr. Rene Jacques
Sex: Male
Age: 36
Passenger Class: Second
Fare: 12.875

Port of Embarkation: Cherbourg
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Died

Using a Decision Tree to Classify Cases

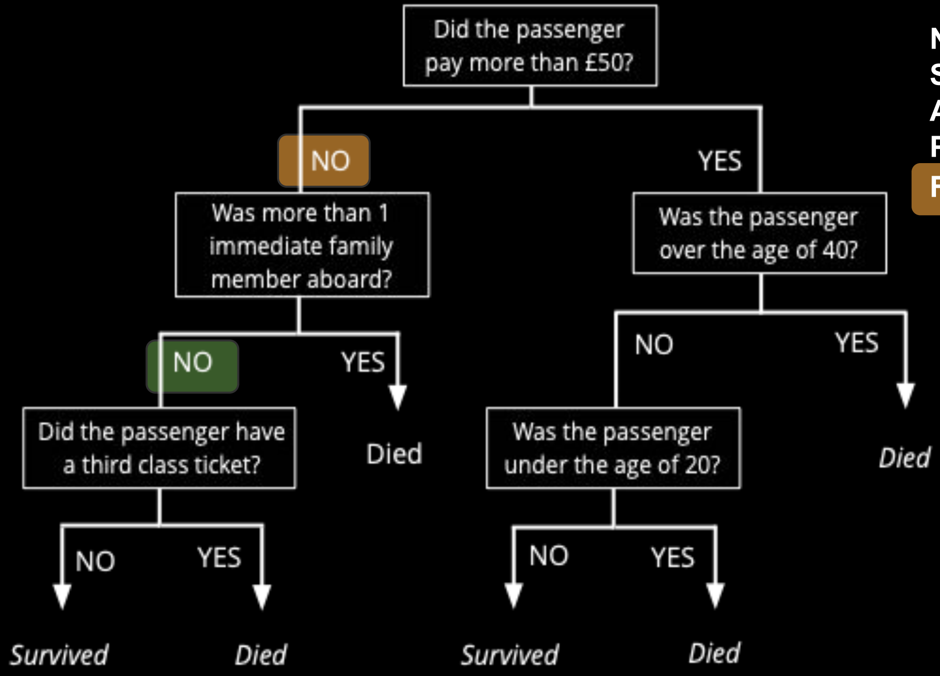


Name: Levy, Mr. Rene Jacques
Sex: Male
Age: 36
Passenger Class: Second
Fare: 12.875

Port of Embarkation: Cherbourg
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Died

Using a Decision Tree to Classify Cases

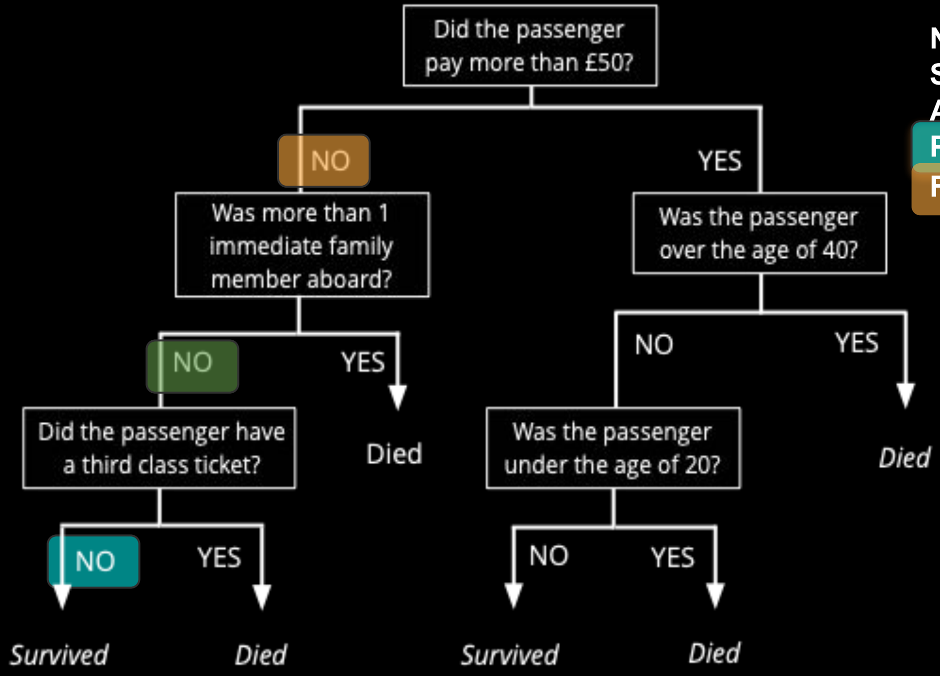


Name: Levy, Mr. Rene Jacques
Sex: Male
Age: 36
Passenger Class: Second
Fare: 12.875

Port of Embarkation: Cherbourg
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Died

Using a Decision Tree to Classify Cases

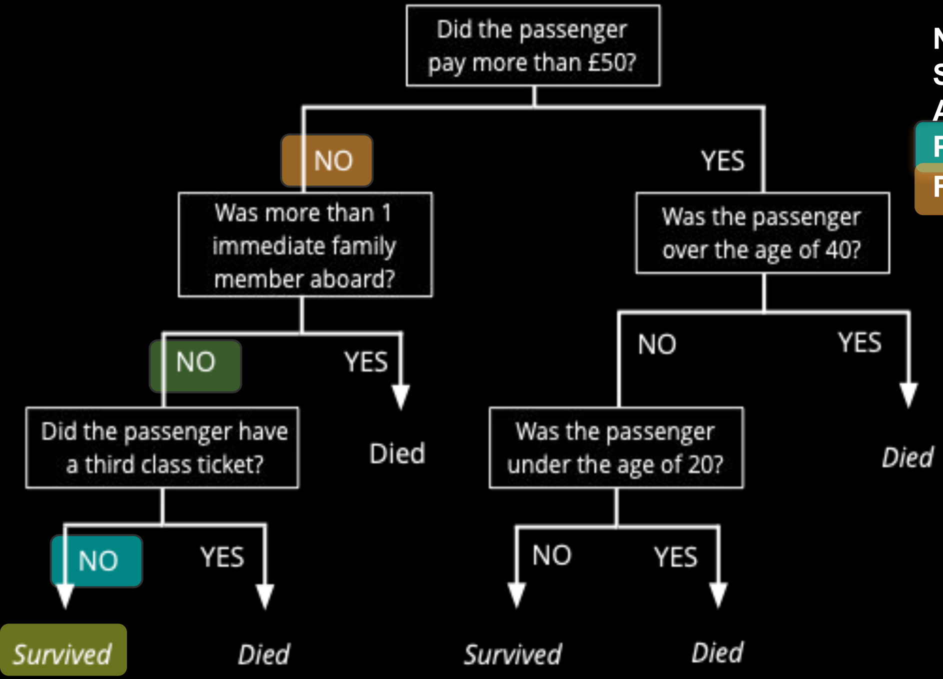


Name: Levy, Mr. Rene Jacques
Sex: Male
Age: 36
Passenger Class: Second
Fare: 12.875

Port of Embarkation: Cherbourg
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Died

Using a Decision Tree to Classify Cases

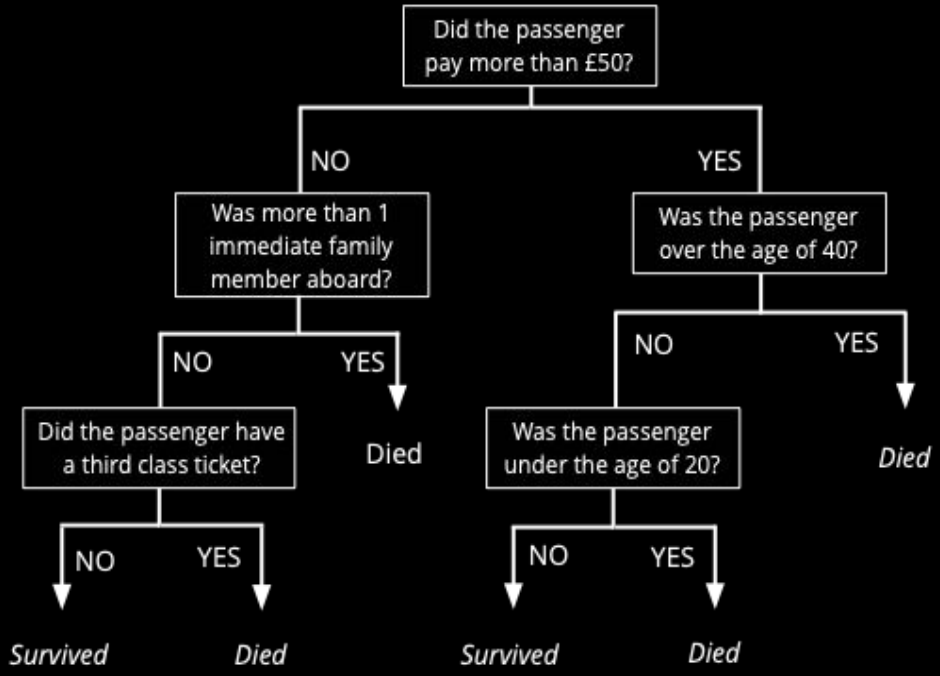


Name: Levy, Mr. Rene Jacques
Sex: Male
Age: 36
Passenger Class: Second
Fare: 12.875

Port of Embarkation: Cherbourg
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Died

Using a Decision Tree to Classify Cases

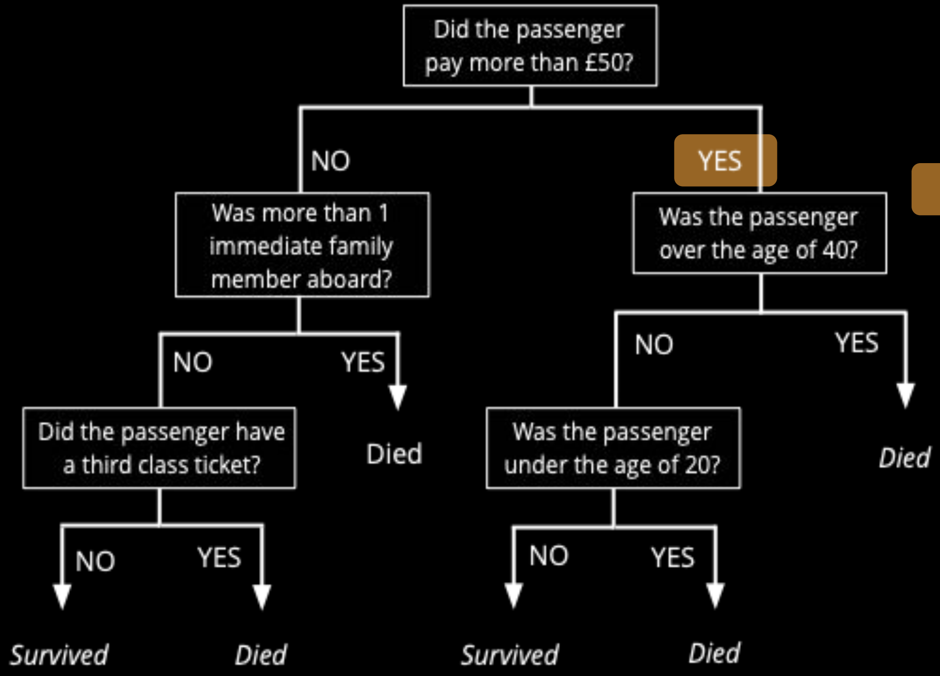


Name: Bing, Mr. Lee
Sex: Male
Age: 32
Passenger Class: Third
Fare: 56.4958

Port of Embarkation: Southampton
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Survived

Using a Decision Tree to Classify Cases

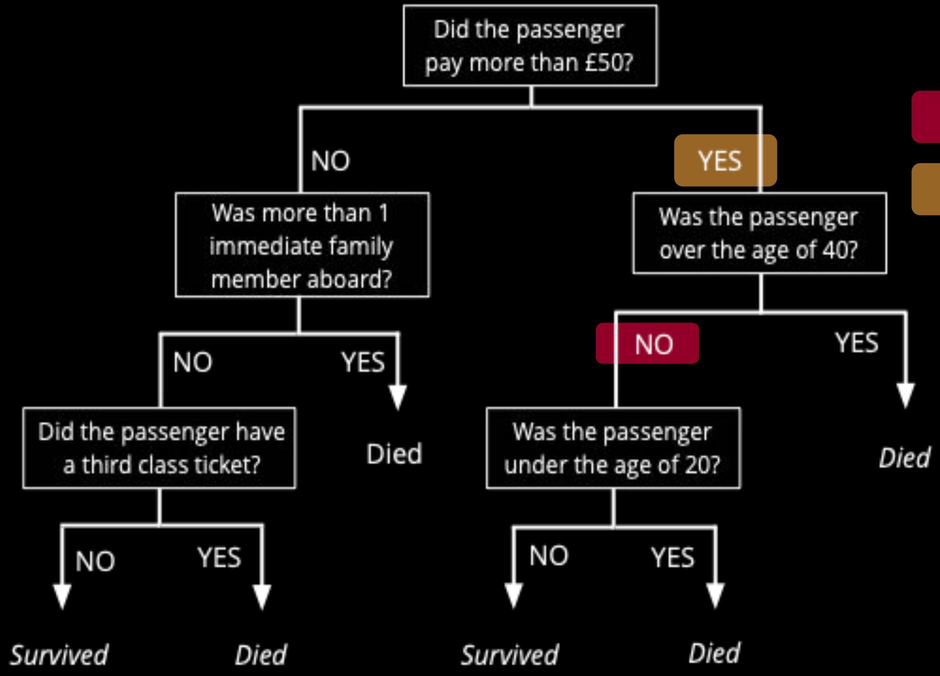


Name: Bing, Mr. Lee
Sex: Male
Age: 32
Passenger Class: Third
Fare: 56.4958

Port of Embarkation: Southampton
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Survived

Using a Decision Tree to Classify Cases

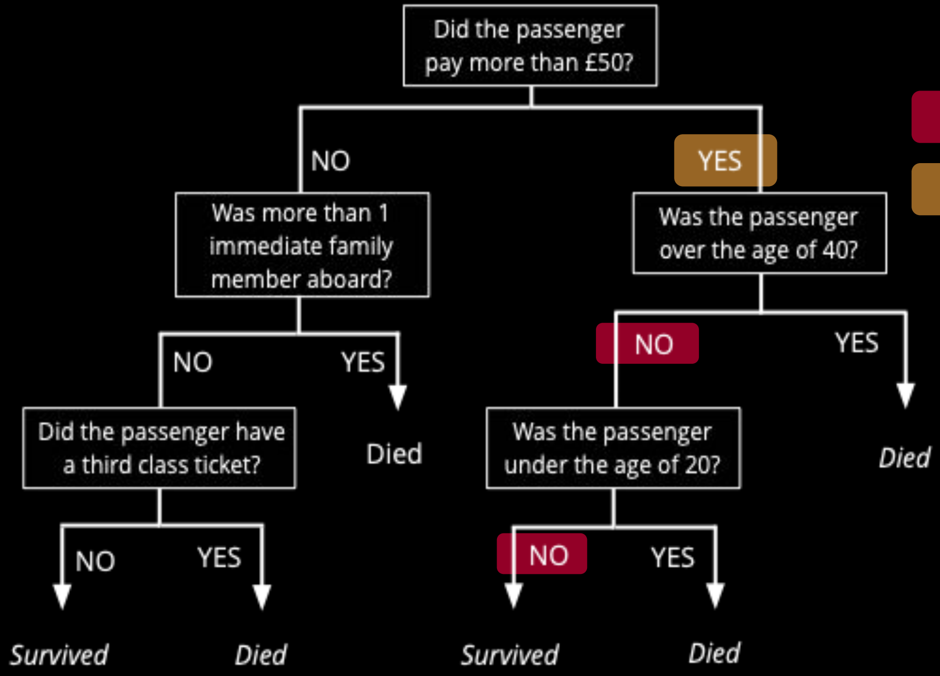


Name: Bing, Mr. Lee
Sex: Male
Age: 32
Passenger Class: Third
Fare: 56.4958

Port of Embarkation: Southampton
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Survived

Using a Decision Tree to Classify Cases

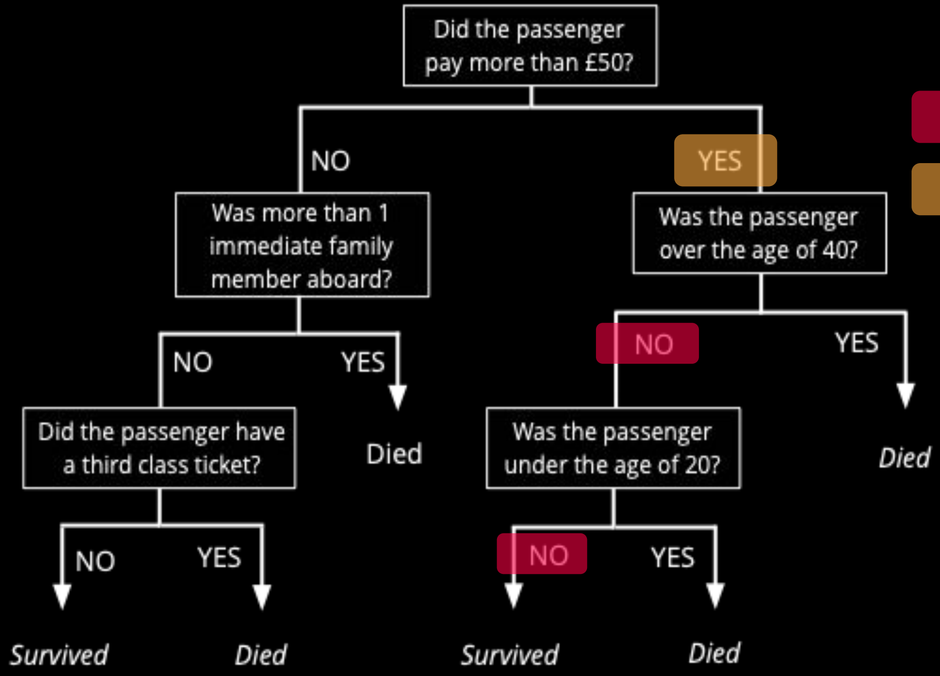


Name: Bing, Mr. Lee
Sex: Male
Age: 32
Passenger Class: Third
Fare: 56.4958

Port of Embarkation: Southampton
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Survived

Using a Decision Tree to Classify Cases

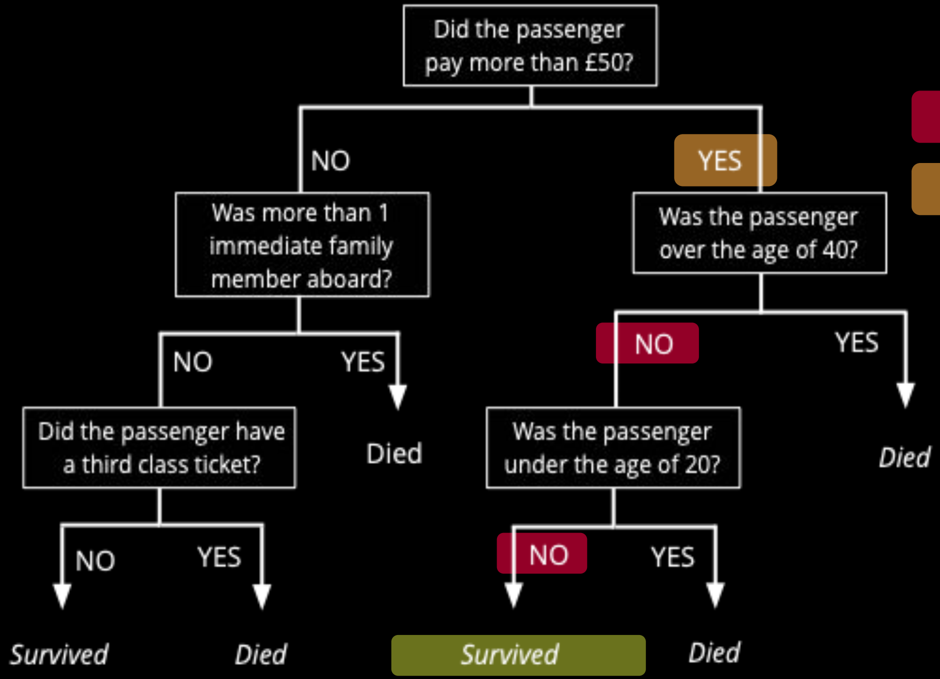


Name: Bing, Mr. Lee
Sex: Male
Age: 32
Passenger Class: Third
Fare: 56.4958

Port of Embarkation: Southampton
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Survived

Using a Decision Tree to Classify Cases



Name: Bing, Mr. Lee
Sex: Male
Age: 32
Passenger Class: Third
Fare: 56.4958

Port of Embarkation: Southampton
Number of Family Members Aboard:
- *Siblings/Spouse:* 0
- *Parents/Children:* 0

Fate: Survived

Evaluating the Model

True Class	Predicted Class	
	Died	Survived
Died		
Survived		

Evaluating the Model

True Class	Predicted Class	
	Died	Survived
Died		Mr. Levy
Survived		

Evaluating the Model

True Class	**Predicted Class**	
	Died	Survived
Died		Mr. Levy
Survived		Mr. Bing

Evaluating the Model

True Class	Predicted Class	
	Died	**Survived**
Died	10 (33.3%)	4 (13.3%)
Survived	6 (20.0%)	10 (33.3%)

Evaluating the Model: Come up with a measure of classification accuracy you could use to evaluate this model.

True Class	Predicted Class: Died	Predicted Class: Survived
Died	10 (33.3%)	4 (13.3%)
Survived	6 (20.0%)	10 (33.3%)

Based on your measure of classification accuracy, how would you determine if this is a “good” model?

True Class	Predicted Class: Died	Predicted Class: Survived
Died	10 (33.3%)	4 (13.3%)
Survived	6 (20.0%)	10 (33.3%)

Based on your measure of classification accuracy, was (this) a "good" model? Explain.

Group B

Group A A model of ~~73~~ 75%. accuracy, while better than chance (50/50), doesn't seem like the best model.

Fairly good since we are erring on the side of the most fatalities.

Group C It was an "ok" model. We would want something in the range of 80-90%