Tourism Insurance

Predicting Unhealthy Days in D.C.

Kate Hannon
Camille Hoff
Tian Liu
Clayton Paulding
Ramin Riahikhoee

Background – D.C. Ducks and AQI

- D.C. Ducks
  - Guided Tours of Washington DC
  - Revenue directly tied to number of tourists– no tourists show up on days with unhealthy weather
  - Insurance to offset revenue loss
  - Blanket *vs.* Selective Weather Insurance

- The Air Quality Index (AQI)
  - AQI > 50 = unhealthy weather

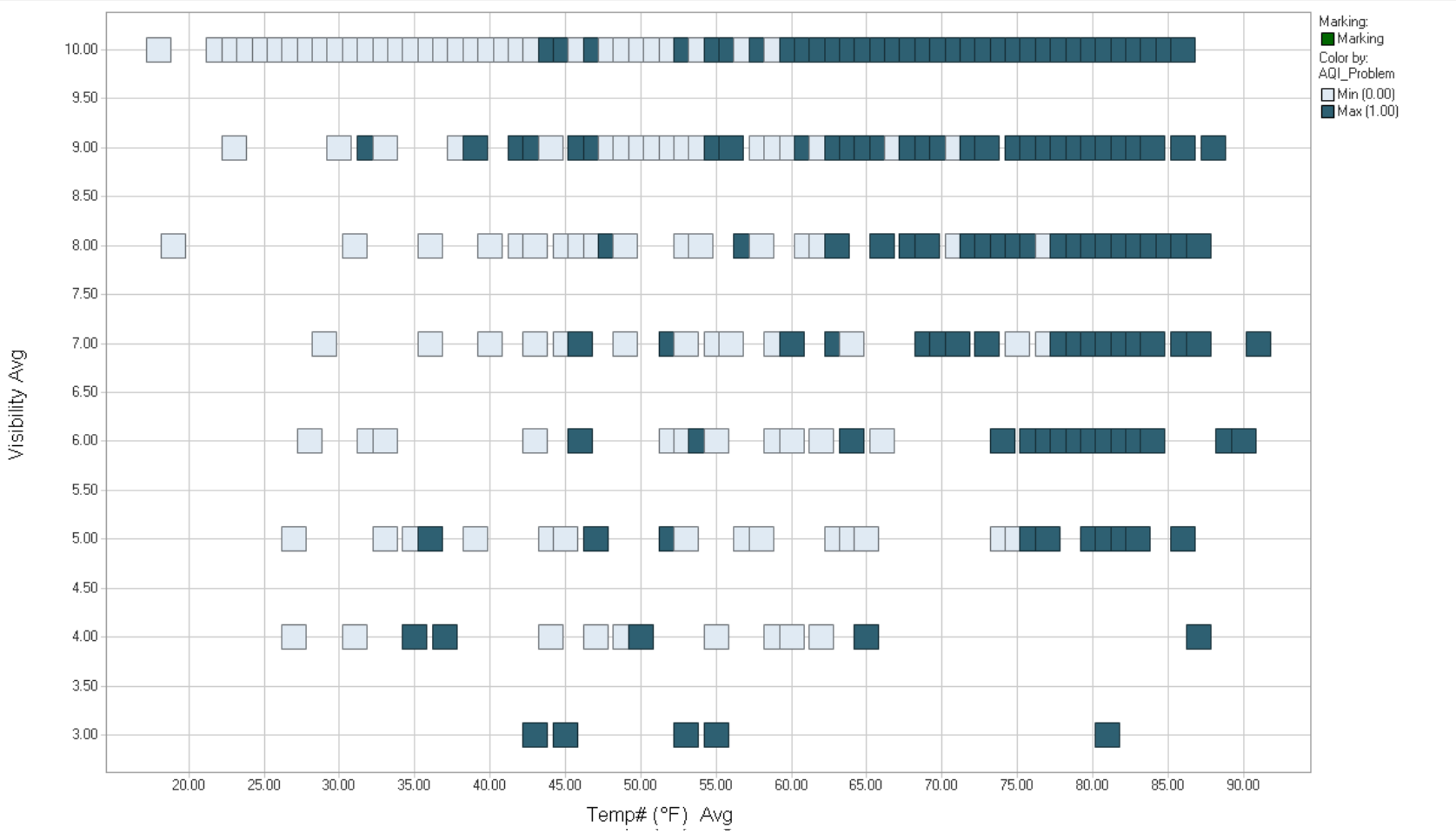
Air Quality Index (AQI) Values	Levels of Health Concern	Colors
When the AQI is in this range:	*...air quality conditions are:*	*...as symbolized by this color:*
0 to 50	Good	Green
51 to 100	Moderate	Yellow
101 to 150	Unhealthy for Sensitive Groups	Orange
151 to 200	Unhealthy	Red
201 to 300	Very Unhealthy	Purple
301 to 500	Hazardous	Maroon

Issue – Minimizing the Cost of Insurance

Temp. (°F) High	Temp. (°F) Avg.	Temp. (°F) Low	Dew Point High	Dew Point Avg.	Dew Point Low	Humidity High	Humidity Avg.	Humidity Low	Sea Level High	Sea Level Avg.	Sea Level Low
Visibility High	Visibility Avg.	Visibility Low	Wind High	Wind Avg.	Gust Speed	Precipitation (inches)	Rain Yes	Tstorm Yes	Fog Yes	Hail Yes	Snow Yes

- Dataset: three years worth of historical weather data
  - April 1, 2005 – March 31, 2008
  - **Important: using Historical Forecasts *vs.* Actual Data**
  - 1,062 records
  - 17 weather-related variables
  - AQI
  - Partitioning to training (2 years) and validation (1 year)

Data Exploration and Processing



Important Variables: Temperature, Wind Speed, Precipitation,

Model

- Model creation
  - Discriminant Analysis, Logistic Regression and Classification Tree
  - Selected: Classification Tree
    - Lowest validation error rate (23.22%)
    - Ease of explanation and use

Validation Data scoring - Summary Report (Using Best Pruned Tree)

Cutoff Prob.Val. for Success (Updatable)	0.5

Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	91	53
0	32	190

Error Report			
Class	# Cases	# Errors	% Error
1	144	53	36.81
0	222	32	14.41
Overall	366	85	23.22

Analysis and Recommendation

Company	Number of Days (in 3 years)	Total Value (3 years)
Number of Days in 3 Year Data Period (1 leap year)	1,096	
Historical Number of High AQI Days in 3 years	379	
Normal AQI days correctly predicted	509	$7,806,155
High AQI days correctly predicted	275	$3,719,600
High AQI days incorrectly predicted as normal	159	($2,177,506)
Normal days incorrectly predicted as High AQI	96	$1,308,013
Cost of blanket insurance, per day	$70	
Cost of selected days insurance, per day	$100	
Cost of blanket insurance (cost per day for 3 years)		$76,720
Cost to insure 379 selected days over 3 years		$37,900
Expected 3-year Profit, blanket insurance		$14,960,400
Expected 3-year Profit, selected days insurance		$14,999,220
Benefit of selected days insurance over blanket insurance		**$38,820**

- If D.C. Ducks can purchase insurance for only those days during the year that it predicts will have an AQI over 50, they can save $38,820 per year.
- Costs to implement this:
  - Coming up with accurate data for such a prediction
  - Developing and using the model

Questions?