**Logistic Regression Variables Lab R Code**

<https://www.ncdc.noaa.gov/cdo-web/>

Pick TAVG, TMAX, TMIN, WSF5, AWND, WT01, WT02, WT03, WT04, WT05, WT11, WT13, WT14, WT15, WT16, WT17

1. Create a season categorical variable
2. Create wind direction categorical variable from the “WDF5” column

While there are multiple correct ways to do this, this is how I did it:

1. Subtracted 23 from each observation in the WDF5 column

RainSeattle2016$NewWindDir<-RainSeattle2016$WDF5-23

1. Created a factor variable from those new values

RainSeattle2016$WindCat<-cut(RainSeattle2016$NewWindDir, c(-22,22,67,112,157,202,247,292,337))

1. Create a factor variable identifying the presence or lack of rain on each day from the “Rain” column
2. Using the glm() function, predict logistic regression models, sequentially, using Season, Average Wind Speed (AWND), Average Temperature (TAVG), highest temperature (TMAX), lowest temperature (TMIN), your created wind direction variable, and highest sustained five minute wind speed (WSF5) in the dataset. Weather Types (see codebook).

Note the AIC for each model.

1. Create residual plots for each variable
2. Identify which variables are significant
3. List the most likely parsimonious model using the information above

PRCP = Precipitation (mm or inches as per user preference, inches to hundredths on Daily Form pdf file)

SNOW = Snowfall (mm or inches as per user preference, inches to tenths on Daily Form pdf file)

SNWD = Snow depth (mm or inches as per user preference, inches on Daily Form pdf file) TMAX = Maximum temperature (Fahrenheit or Celsius as per user preference, Fahrenheit to tenths on Daily Form pdf file

TMIN = Minimum temperature (Fahrenheit or Celsius as per user preference, Fahrenheit to tenths on Daily Form

The other values are:

ACMC = Average cloudiness midnight to midnight from 30-second ceilometer data (percent) ACMH = Average cloudiness midnight to midnight from manual observations (percent) ACSC = Average cloudiness sunrise to sunset from 30-second ceilometer data (percent) ACSH = Average cloudiness sunrise to sunset from manual observations (percent)

AWND = Average daily wind speed (meters per second or miles per hour as per user preference)

DAEV = Number of days included in the multiday evaporation total (MDEV)

DAPR = Number of days included in the multiday precipitation total (MDPR)

DASF = Number of days included in the multiday snowfall total (MDSF)

DATN = Number of days included in the multiday minimum temperature (MDTN)

DATX = Number of days included in the multiday maximum temperature (MDTX)

DAWM = Number of days included in the multiday wind movement (MDWM)

DWPR = Number of days with non-zero precipitation included in multiday precipitation total (MDPR)

EVAP = Evaporation of water from evaporation pan (mm or inches as per user preference, or hundredths of inches on Daily Form pdf file)

FMTM = Time of fastest mile or fastest 1-minute wind (hours and minutes, i.e., HHMM) FRGB = Base of frozen ground layer (cm or inches as per user preference)

FRGT = Top of frozen ground layer (cm or inches as per user preference)

FRTH = Thickness of frozen ground layer (cm or inches as per user preference)

GAHT = Difference between river and gauge height (cm or inches as per user preference) MDEV = Multiday evaporation total (mm or inches as per user preference; use with DAEV) MDPR = Multiday precipitation total (mm or inches as per user preference; use with DAPR and DWPR, if available)

MDSF = Multiday snowfall total (mm or inches as per user preference)

MDTN = Multiday minimum temperature (Fahrenheit or Celsius as per user preference ; use with DATN)

MDTX = Multiday maximum temperature (Fahrenheit or Celsius as per user preference ; use with DATX)

MDWM = Multiday wind movement (miles or km as per user preference)

MNPN = Daily minimum temperature of water in an evaporation pan (Fahrenheit or Celsius as per user preference)

MXPN = Daily maximum temperature of water in an evaporation pan (Fahrenheit or Celsius as per user preference)

PGTM = Peak gust time (hours and minutes, i.e., HHMM)

PSUN = Daily percent of possible sunshine (percent)

SN\*# = Minimum soil temperature where \* corresponds to a code for ground cover and # corresponds to a code for soil depth (Fahrenheit or Celsius as per user preference) Ground cover codes include the following: 0 = unknown 1 = grass 2 = fallow 3 = bare ground 4 = brome grass 5 = sod 6 = straw mulch 7 = grass muck 8 = bare muck Depth codes include the following: 1 = 5 cm 2 = 10 cm 3 = 20 cm 4 = 50 cm 5 = 100 cm 6 = 150 cm 7 = 180 cm SX\*# = Maximum soil temperature where \* corresponds to a code for ground cover and # corresponds to a code for soil depth. See SN\*# for depth codes. (Fahrenheit or Celsius as per user preference) THIC = Thickness of ice on water (inches or mm as per user preference) TOBS = Temperature at the time of observation (Fahrenheit or Celsius as per user preference) TSUN = Daily total sunshine (minutes) WDF1 = Direction of fastest 1-minute wind (degrees) WDF2 = Direction of fastest 2-minute wind (degrees)

WDF5 = Direction of fastest 5-second wind (degrees)

WDFG = Direction of peak wind gust (degrees)

WDFI = Direction of highest instantaneous wind (degrees)

WDFM = Fastest mile wind direction (degrees)

WDMV = 24-hour wind movement (km or miles as per user preference, miles on Daily Form pdf file)

WESD = Water equivalent of snow on the ground (inches or mm as per user preference) WESF = Water equivalent of snowfall (inches or mm as per user preference)

WSF1 = Fastest 1-minute wind speed (miles per hour or meters per second as per user preference)

WSF2 = Fastest 2-minute wind speed (miles per hour or meters per second as per user preference)

WSF5 = Fastest 5-second wind speed (miles per hour or meters per second as per user preference)

WSFG = Peak guest wind speed (miles per hour or meters per second as per user preference) WSFI = Highest instantaneous wind speed (miles per hour or meters per second as per user preference) WSFM = Fastest mile wind speed (miles per hour or meters per second as per user preference)

WT\*\* = Weather Type where \*\* has one of the following values: 01 = Fog, ice fog, or freezing fog (may include heavy fog) 02 = Heavy fog or heaving freezing fog (not always distinguished from fog) 03 = Thunder 04 = Ice pellets, sleet, snow pellets, or small hail 05 = Hail (may include small hail) 06 = Glaze or rime 07 = Dust, volcanic ash, blowing dust, blowing sand, or blowing obstruction 08 = Smoke or haze 09 = Blowing or drifting snow 10 = Tornado, waterspout, or funnel cloud 11 = High or damaging winds 12 = Blowing spray 13 = Mist 14 = Drizzle 15 = Freezing drizzle 16 = Rain (may include freezing rain, drizzle, and freezing drizzle) 17 = Freezing rain 18 = Snow, snow pellets, snow grains, or ice crystals 19 = Unknown source of precipitation 21 = Ground fog 22 = Ice fog or freezing fog WVxx = Weather in the Vicinity where “xx” has one of the following values 01 = Fog, ice fog, or freezing fog (may include heavy fog) 03 = Thunder 07 = Ash, dust, sand, or other blowing obstruction 18 = Snow or ice crystals 20 = Rain or snow shower

WT01

WT02

WT03

WT04

WT05

WT11

WT12

WT13

WT14

WT15

WT16

WT17