Economic Output and Risk Factors

Nathan Seto

What is the relationship between health risk concentrations and economic output?

Calculate correlations between environmental exposures and farm sales (and/or other econ. variables) to understand economic linkages.

To see coefficients and significance, follow this link: Correlation Shiny.
The values used as variables of economic output are listed below. The predictor variables are all the variables in each data set that are not in the following list of variables. Variables ending in "$_operation" denote "dollars per operation".

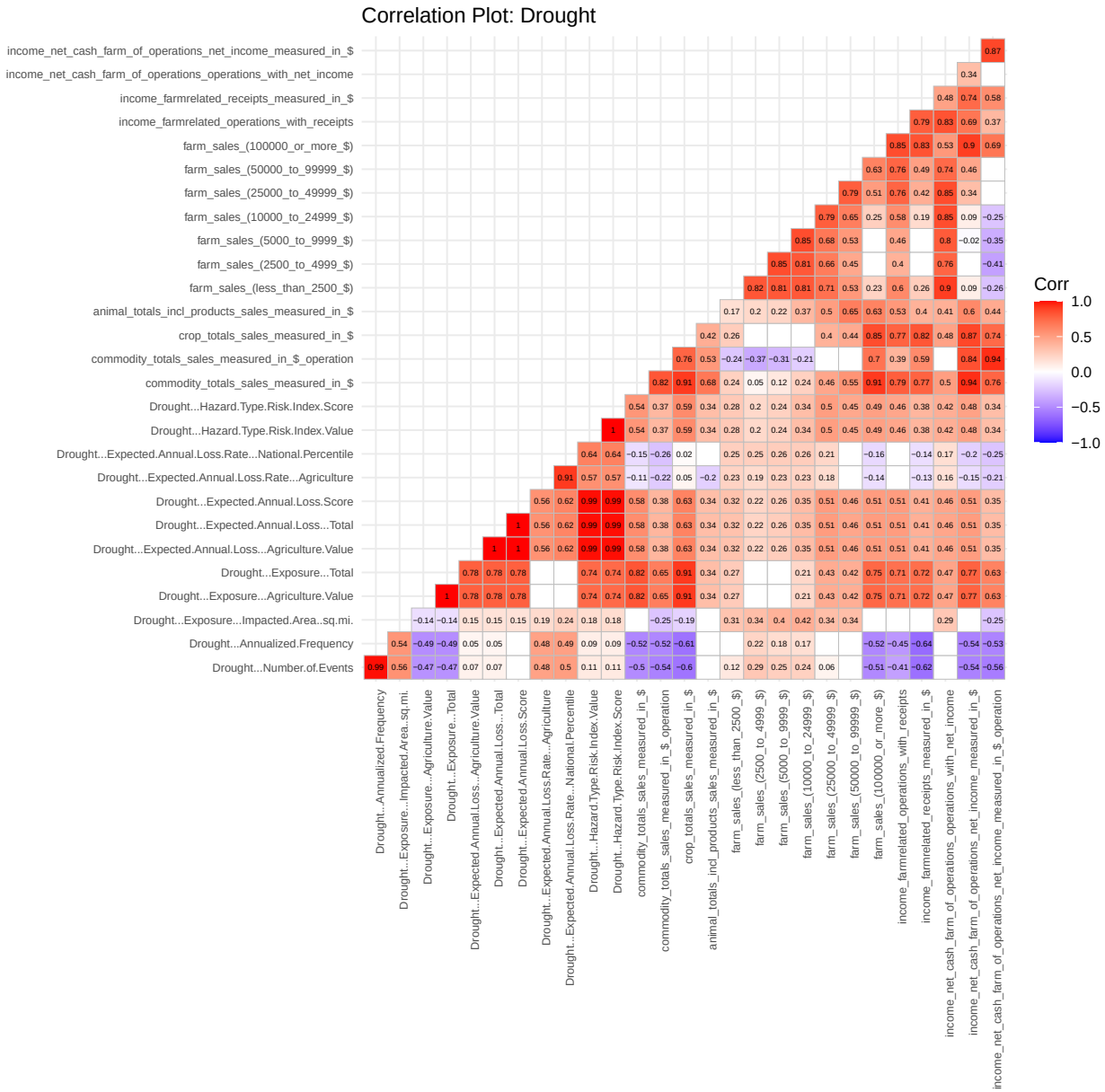
```
## [[1]]
##  [1] "commodity_totals_sales_measured_in_$"
##  [2] "commodity_totals_sales_measured_in_$_operation"
##  [3] "crop_totals_sales_measured_in_$"
##  [4] "animal_totals_incl_products_sales_measured_in_$"
##  [5] "commodity_totals_operations_with_sales"
##  [6] "farm_sales_(less_than_2500_$)"
##  [7] "farm_sales_(2500_to_4999_$)"
##  [8] "farm_sales_(5000_to_9999_$)"
##  [9] "farm_sales_(10000_to_24999_$)"
## [10] "farm_sales_(25000_to_49999_$)"
## [11] "farm_sales_(50000_to_99999_$)"
## [12] "farm_sales_(100000_or_more_$)"
## [13] "income_farmrelated_operations_with_receipts"
## [14] "income_farmrelated_receipts_measured_in_$"
## [15] "income_net_cash_farm_of_operations_operations_with_net_income"
## [16] "income_net_cash_farm_of_operations_net_income_measured_in_$"
## [17] "income_net_cash_farm_of_operations_net_income_measured_in_$_operation"
```

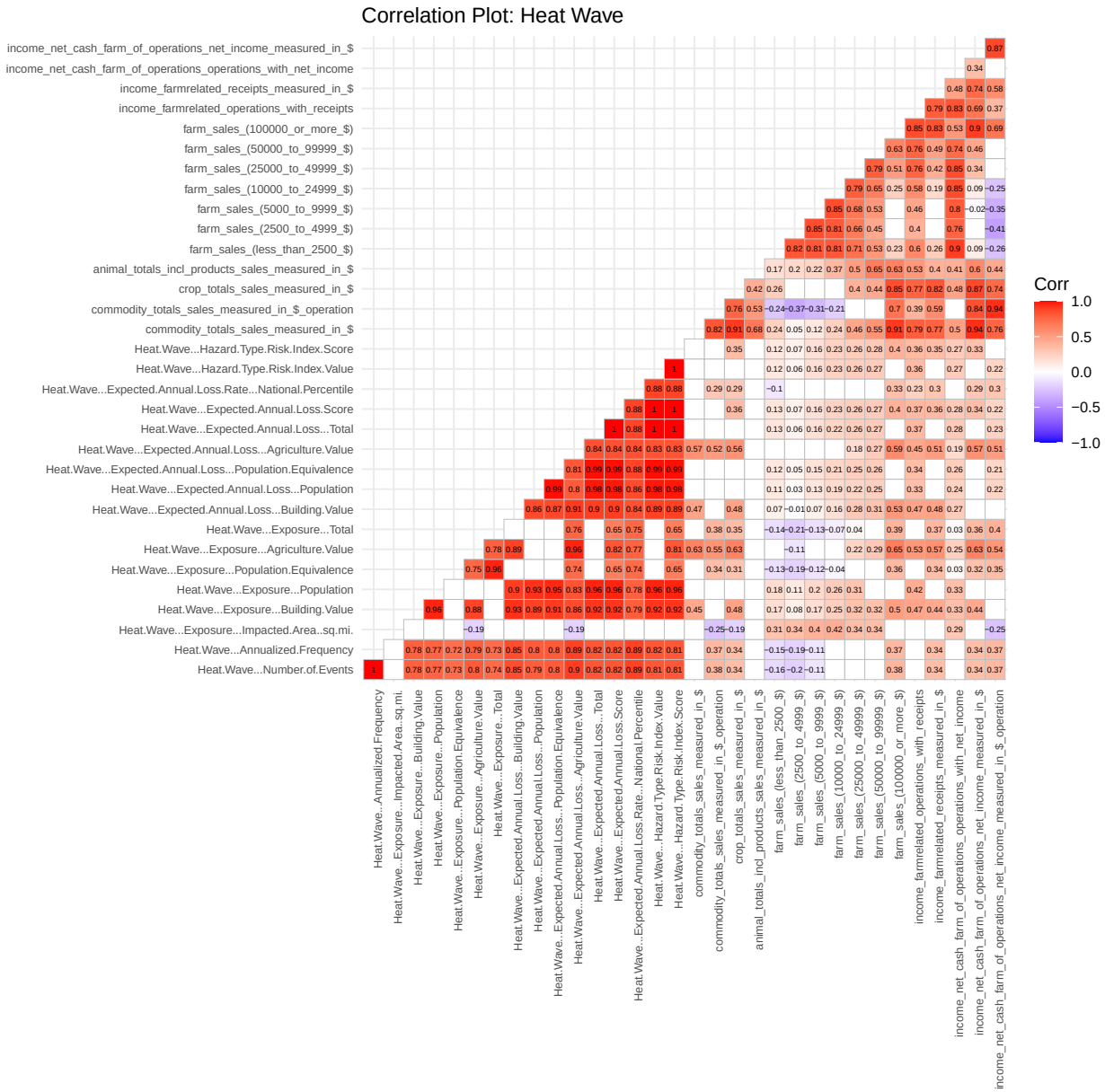
Methods
I started this process by collecting separate data sets for all counties in the states covered by HICAHS from the United States Department of Agriculture agricultural census database. I then merged all these data sets into a single file. Since the variables were in rows, I pivoted the data set so the variables were made into columns. Following this, I merged the data containing natural disaster/weather variables, creating the complete data I desired to work with. See 'Processing/output_Risk.R' for specific methods. Correlation coefficients were reached by taking one predictor and calculating the correlation coefficient and p-value between the remaining variables. This was done for all 17 response variables above. These values were then entered into a table to be displayed. To create the correlation plots found below, the original data set was split into 12 data sets by the type of natural disaster/weather factor, then correlations were calculated for each pair and graphed. The 17 response variables are identical throughout all 12 data sets. All work was completed using R and RStudio.

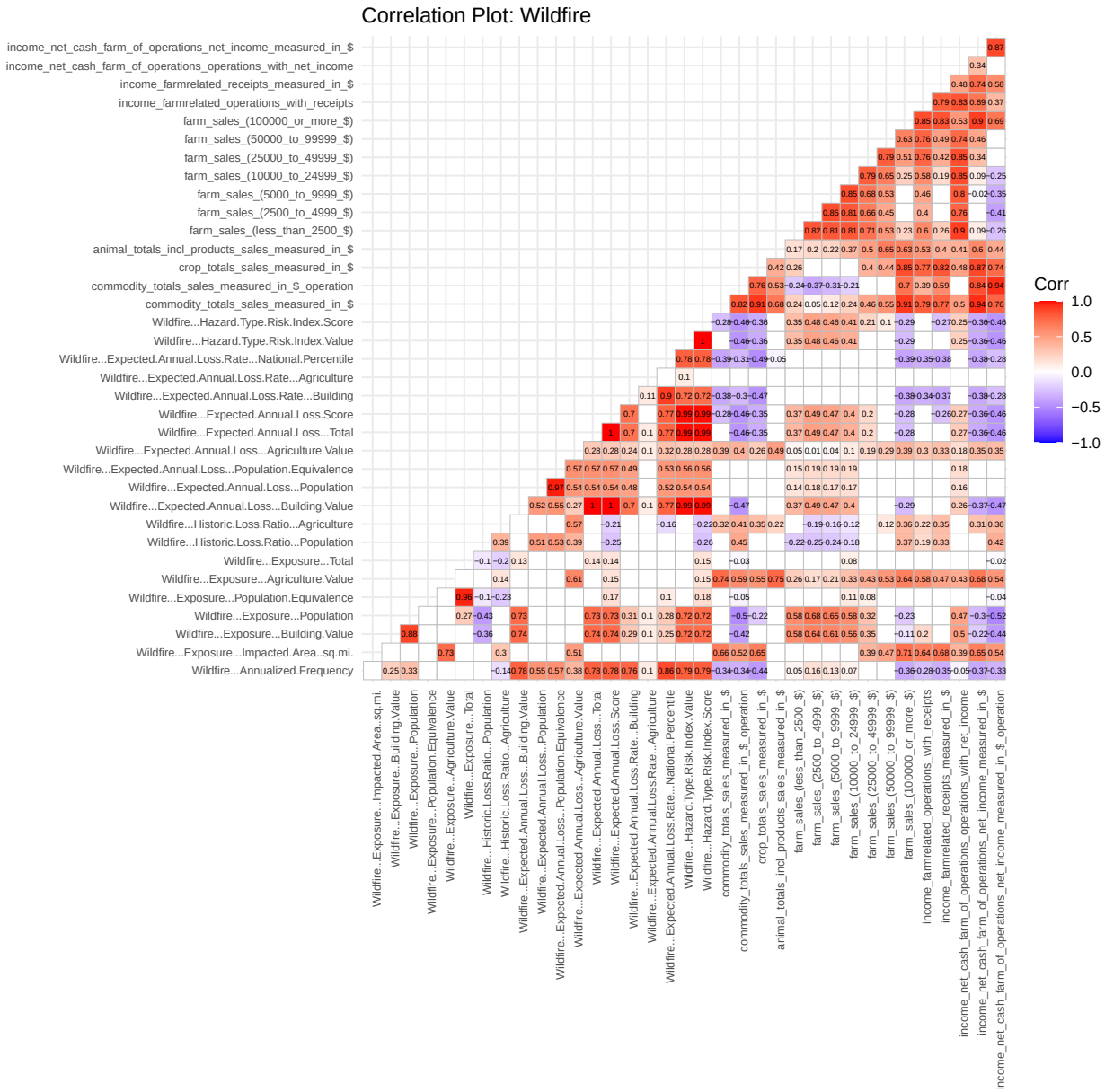
The Correlations
Most of the correlation coefficients are between 0.00 and 0.30. The higher the correlation coefficient is, the smaller the proportion within the data. It appears that the range of 0.60 to 0.80 has some of the smallest (non-zero) proportions of p-values between all 12 data sets. Looking at the entire volume, variables of economic output and natural disasters/weather have generally weak correlations, whether positive or negative. One would expect that natural disasters and poor weather would correlate heavily with lesser output, but the current data set fails to capture this alternative hypothesis as well as expected. This may be because the data is not associated with location-specific data, but only counties and states. In other words, a wildfire or a tornado that does not occur near any farms will have an extremely small effect on its output, if any effect at all. A deeper investigation into locations where natural disasters and poor weather conditions have occurred near farms would be more informative, but this is more costly and complicated. Correlations using location data, such as coordinates, would be the easiest way to do this.

Significance
Many p-values in the table are shown as zero. This means either the p-value is, in fact, zero, or very nearly zero. P-values equal to zero mean that the result is statistically impossible to occur under the null hypothesis. The number of p-values that are shown as zero is alarming, which brings to question the methods of testing I used or the structure of the data. However, I did use 'cor.test()' whose results I have no reason to question the validity of. Furthermore, I cannot remember the last time I heard of a tornado touching down in Colorado (if it has happened in my lifetime at all), avalanches are extremely unlikely to occur near farms and affect their output since locations of either are usually uncorrelated. Simply put, some locations experience different weather risks at different levels. It is for these reasons the p-values are as valid as possible.

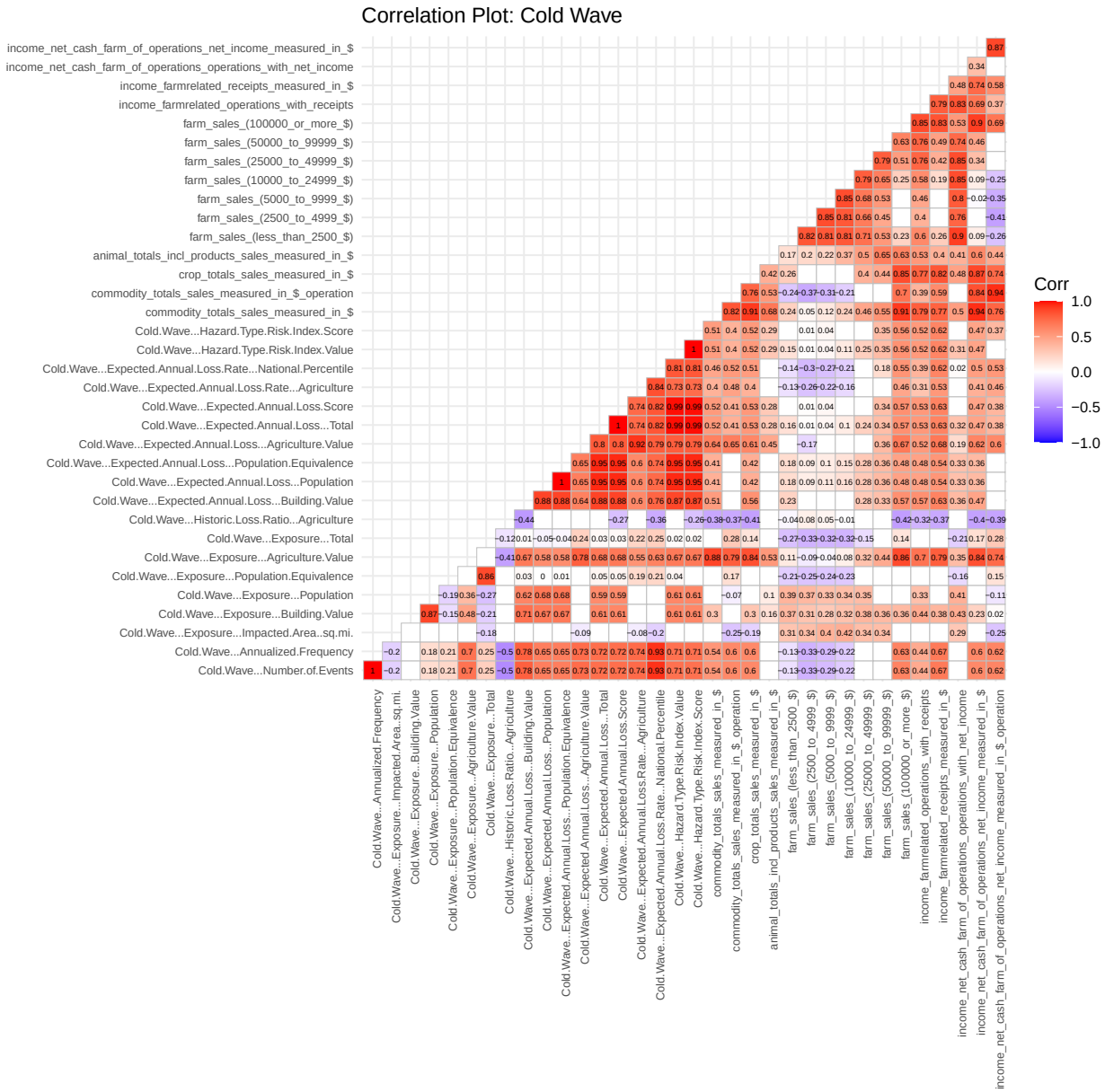
Correlograms
Below are a series of correlation plots between variables of economic output and different weather/natural disaster predictors. The main focus is the rectangle towards the center of the plots, more prominent in some plots than others. The disaster/weather variables are naturally correlated with each other a majority of the time. The response variables are naturally correlated with each other to some degree. This is to be expected.

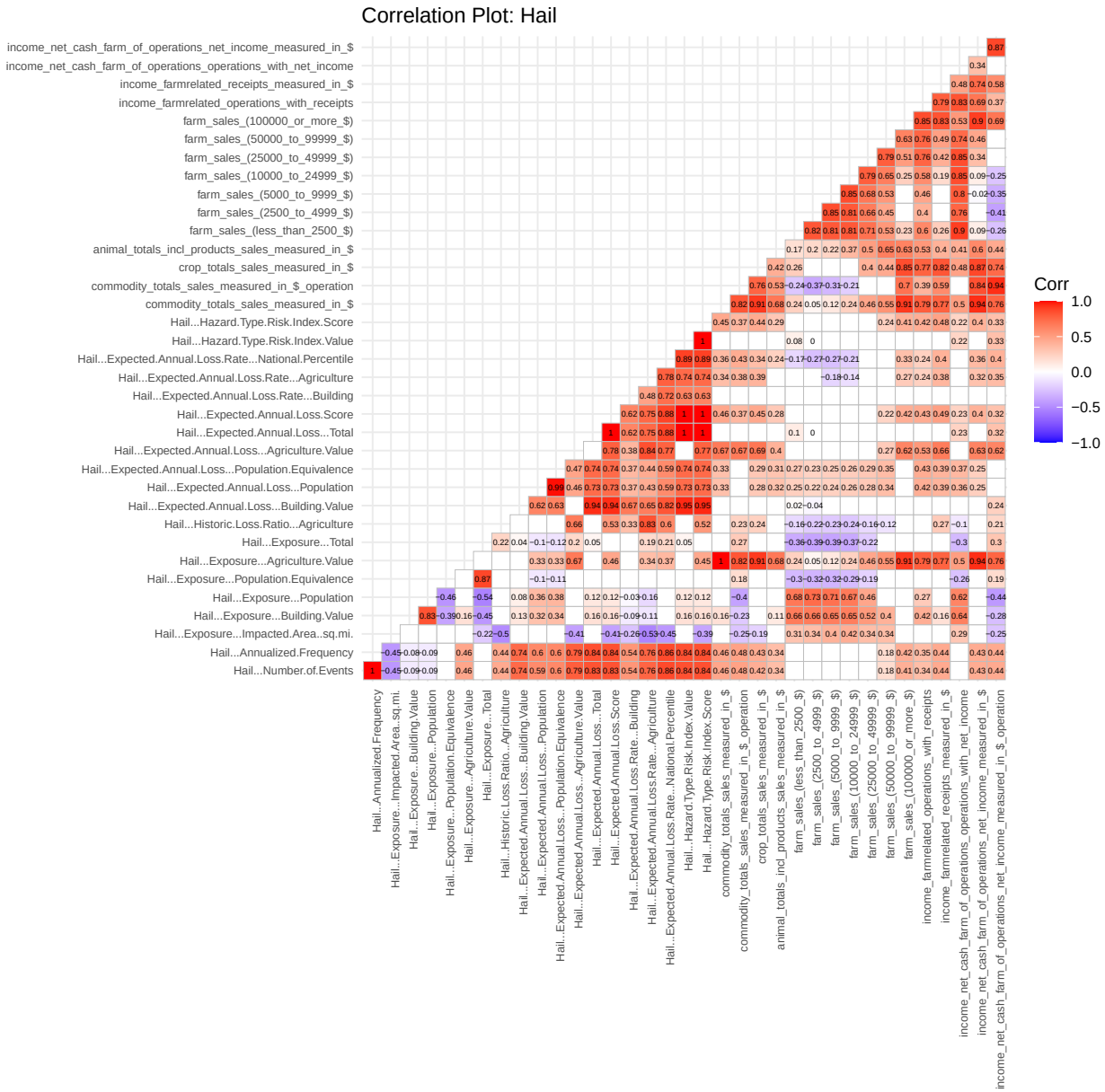

Correlation Plot: Drought
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_farmrelated_receipts_measured_in_$
income_farmrelated_operations_with_receipts
farm_sales_(100000_or_more_$)
farm_sales_(50000_to_99999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(2500_to_4999_$)
farm_sales_(less_than_2500_$)
animal_totals_incl_products_sales_measured_in_$
crop_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
commodity_totals_sales_measured_in_$
Drought...Hazard.Type.Risk.Index.Score
Drought...Hazard.Type.Risk.Index.Value
Drought...Expected.Annual.Loss.Rate...National.Percentile
Drought...Expected.Annual.Loss.Rate...Agriculture
Drought...Expected.Annual.Loss.Score
Drought...Expected.Annual.Loss...Total
Drought...Expected.Annual.Loss...Agriculture.Value
Drought...Exposure...Total
Drought...Exposure...Agriculture.Value
Drought...Exposure...Impacted.Area..sq.mi.
Drought...Annualized.Frequency
Drought...Number.of.Events
Drought...Annualized.Frequency
Drought...Exposure...Impacted.Area..sq.mi.
Drought...Exposure...Agriculture.Value
Drought...Exposure...Total
Drought...Expected.Annual.Loss...Agriculture.Value
Drought...Expected.Annual.Loss...Total
Drought...Expected.Annual.Loss.Score
Drought...Expected.Annual.Loss.Rate...Agriculture
Drought...Expected.Annual.Loss.Rate...National.Percentile
Drought...Hazard.Type.Risk.Index.Value
Drought...Hazard.Type.Risk.Index.Score
commodity_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
crop_totals_sales_measured_in_$
animal_totals_incl_products_sales_measured_in_$
farm_sales_(less_than_2500_$)
farm_sales_(2500_to_4999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(50000_to_99999_$)
farm_sales_(100000_or_more_$)
income_farmrelated_operations_with_receipts
income_farmrelated_receipts_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_net_income_measured_in_$_operation
Corr
1.0
0.5
0.0
−0.5
−1.0

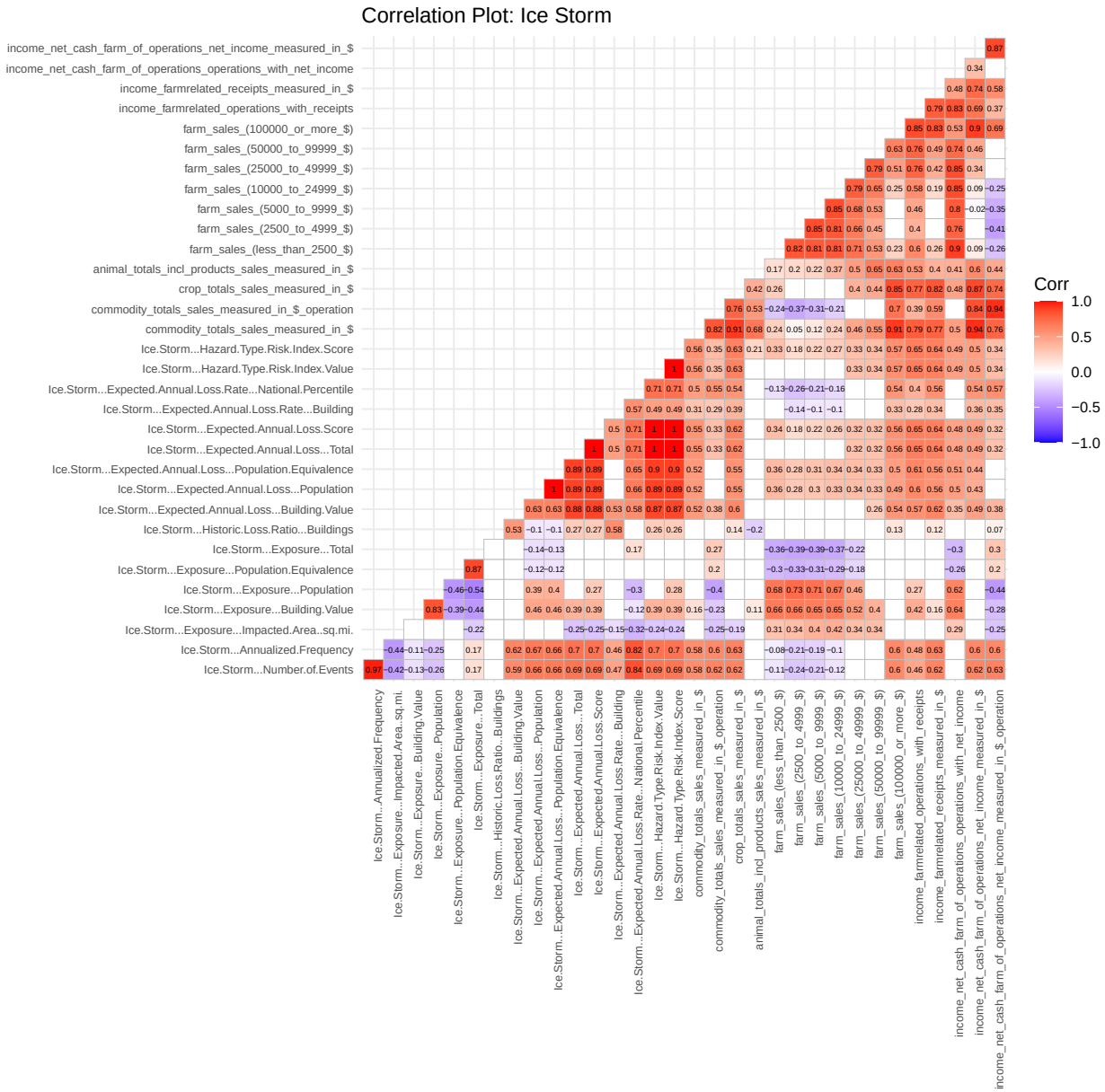

Correlation Plot: Heat Wave
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_farmrelated_receipts_measured_in_$
income_farmrelated_operations_with_receipts
farm_sales_(100000_or_more_$)
farm_sales_(50000_to_99999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(2500_to_4999_$)
farm_sales_(less_than_2500_$)
animal_totals_incl_products_sales_measured_in_$
crop_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
commodity_totals_sales_measured_in_$
Heat.Wave...Hazard.Type.Risk.Index.Score
Heat.Wave...Hazard.Type.Risk.Index.Value
Heat.Wave...Expected.Annual.Loss.Rate...National.Percentile
Heat.Wave...Expected.Annual.Loss.Score
Heat.Wave...Expected.Annual.Loss...Total
Heat.Wave...Expected.Annual.Loss...Agriculture.Value
Heat.Wave...Expected.Annual.Loss...Population.Equivalence
Heat.Wave...Expected.Annual.Loss...Population
Heat.Wave...Expected.Annual.Loss...Building.Value
Heat.Wave...Exposure...Total
Heat.Wave...Exposure...Agriculture.Value
Heat.Wave...Exposure...Population.Equivalence
Heat.Wave...Exposure...Population
Heat.Wave...Exposure...Building.Value
Heat.Wave...Exposure...Impacted.Area..sq.mi.
Heat.Wave...Annualized.Frequency
Heat.Wave...Number.of.Events
Corr
1.0
0.5
0.0
−0.5
−1.0


Correlation Plot: Wildfire
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_farmrelated_receipts_measured_in_$
income_farmrelated_operations_with_receipts
farm_sales_(100000_or_more_$)
farm_sales_(50000_to_99999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(2500_to_4999_$)
farm_sales_(less_than_2500_$)
animal_totals_incl_products_sales_measured_in_$
crop_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
commodity_totals_sales_measured_in_$
Wildfire...Hazard.Type.Risk.Index.Score
Wildfire...Hazard.Type.Risk.Index.Value
Wildfire...Expected.Annual.Loss.Rate...National.Percentile
Wildfire...Expected.Annual.Loss.Rate...Agriculture
Wildfire...Expected.Annual.Loss.Rate...Building
Wildfire...Expected.Annual.Loss.Score
Wildfire...Expected.Annual.Loss...Total
Wildfire...Expected.Annual.Loss...Agriculture.Value
Wildfire...Expected.Annual.Loss...Population.Equivalence
Wildfire...Expected.Annual.Loss...Population
Wildfire...Expected.Annual.Loss...Building.Value
Wildfire...Historic.Loss.Ratio...Agriculture
Wildfire...Historic.Loss.Ratio...Population
Wildfire...Exposure...Total
Wildfire...Exposure...Agriculture.Value
Wildfire...Exposure...Population.Equivalence
Wildfire...Exposure...Population
Wildfire...Exposure...Building.Value
Wildfire...Exposure...Impacted.Area..sq.mi.
Wildfire...Annualized.Frequency
Corr
1.0
0.5
0.0
−0.5
−1.0

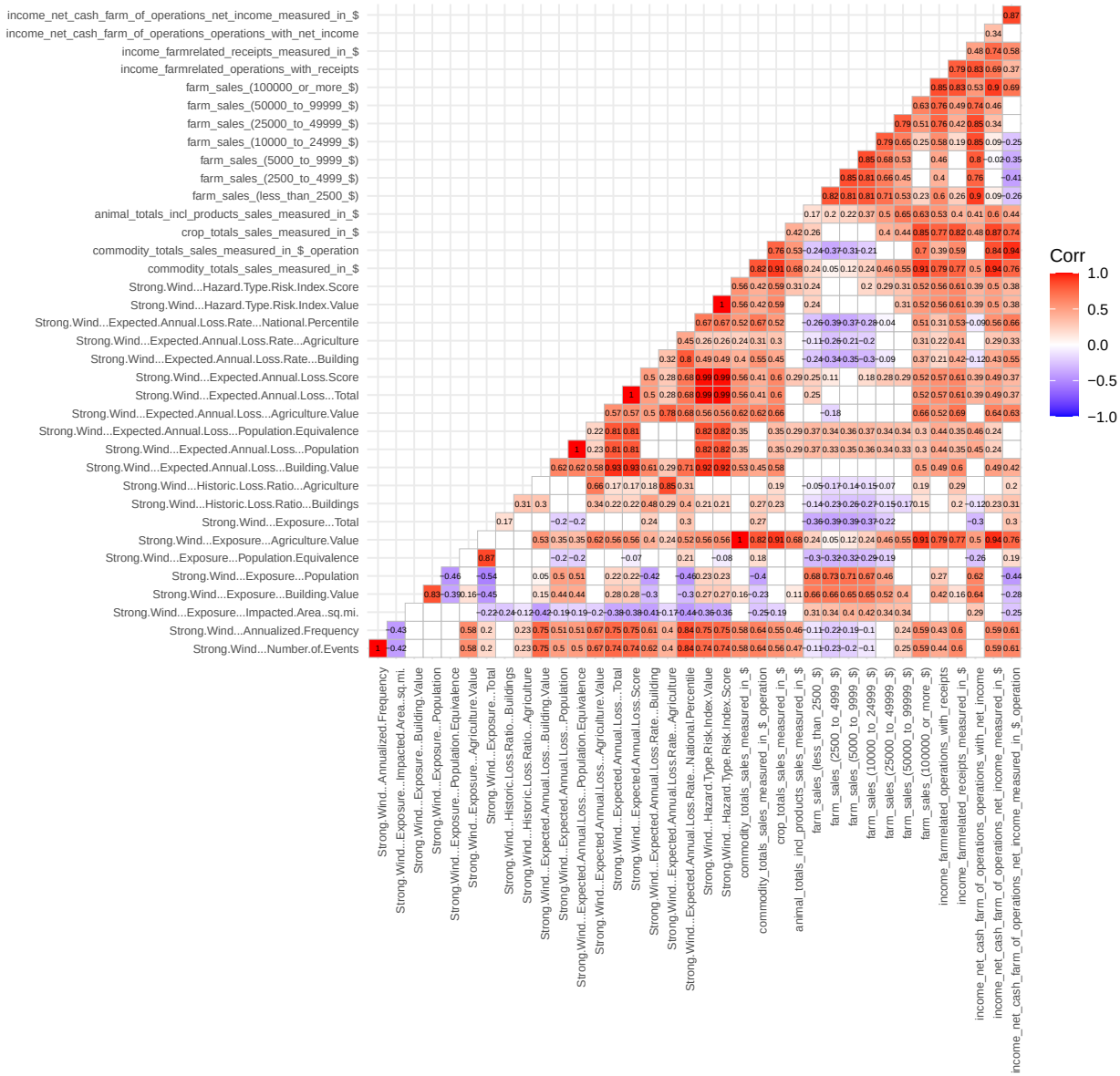
Correlation Plot: Winter Weather


Correlation Plot: Cold Wave
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_farmrelated_receipts_measured_in_$
income_farmrelated_operations_with_receipts
farm_sales_(100000_or_more_$)
farm_sales_(50000_to_99999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(2500_to_4999_$)
farm_sales_(less_than_2500_$)
animal_totals_incl_products_sales_measured_in_$
crop_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
commodity_totals_sales_measured_in_$
Cold.Wave...Hazard.Type.Risk.Index.Score
Cold.Wave...Hazard.Type.Risk.Index.Value
Cold.Wave...Expected.Annual.Loss.Rate...National.Percentile
Cold.Wave...Expected.Annual.Loss.Rate...Agriculture
Cold.Wave...Expected.Annual.Loss.Score
Cold.Wave...Expected.Annual.Loss...Total
Cold.Wave...Expected.Annual.Loss...Agriculture.Value
Cold.Wave...Expected.Annual.Loss...Population.Equivalence
Cold.Wave...Expected.Annual.Loss...Population
Cold.Wave...Expected.Annual.Loss...Building.Value
Cold.Wave...Historic.Loss.Ratio...Agriculture
Cold.Wave...Exposure...Total
Cold.Wave...Exposure...Agriculture.Value
Cold.Wave...Exposure...Population.Equivalence
Cold.Wave...Exposure...Population
Cold.Wave...Exposure...Building.Value
Cold.Wave...Exposure...Impacted.Area..sq.mi.
Cold.Wave...Annualized.Frequency
Cold.Wave...Number.of.Events
Corr
1.0
0.5
0.0
−0.5
−1.0

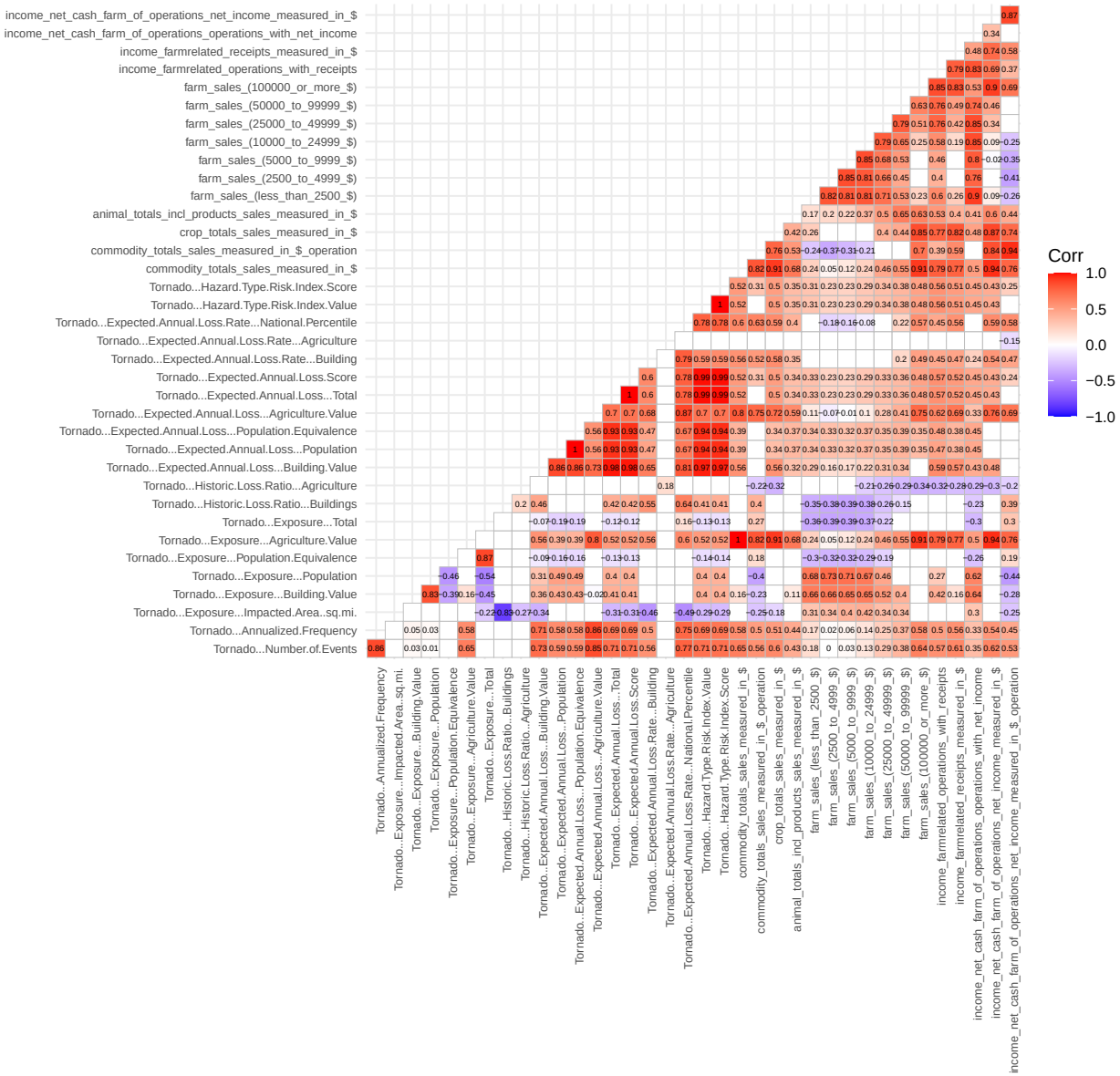

Correlation Plot: Hail
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_farmrelated_receipts_measured_in_$
income_farmrelated_operations_with_receipts
farm_sales_(100000_or_more_$)
farm_sales_(50000_to_99999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(2500_to_4999_$)
farm_sales_(less_than_2500_$)
animal_totals_incl_products_sales_measured_in_$
crop_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
commodity_totals_sales_measured_in_$
Hail...Hazard.Type.Risk.Index.Score
Hail...Hazard.Type.Risk.Index.Value
Hail...Expected.Annual.Loss.Rate...National.Percentile
Hail...Expected.Annual.Loss.Rate...Agriculture
Hail...Expected.Annual.Loss.Rate...Building
Hail...Expected.Annual.Loss.Score
Hail...Expected.Annual.Loss...Total
Hail...Expected.Annual.Loss...Agriculture.Value
Hail...Expected.Annual.Loss...Population.Equivalence
Hail...Expected.Annual.Loss...Population
Hail...Expected.Annual.Loss...Building.Value
Hail...Historic.Loss.Ratio...Agriculture
Hail...Exposure...Total
Hail...Exposure...Agriculture.Value
Hail...Exposure...Population.Equivalence
Hail...Exposure...Population
Hail...Exposure...Building.Value
Hail...Exposure...Impacted.Area..sq.mi.
Hail...Annualized.Frequency
Hail...Number.of.Events
Corr
1.0
0.5
0.0
−0.5
−1.0

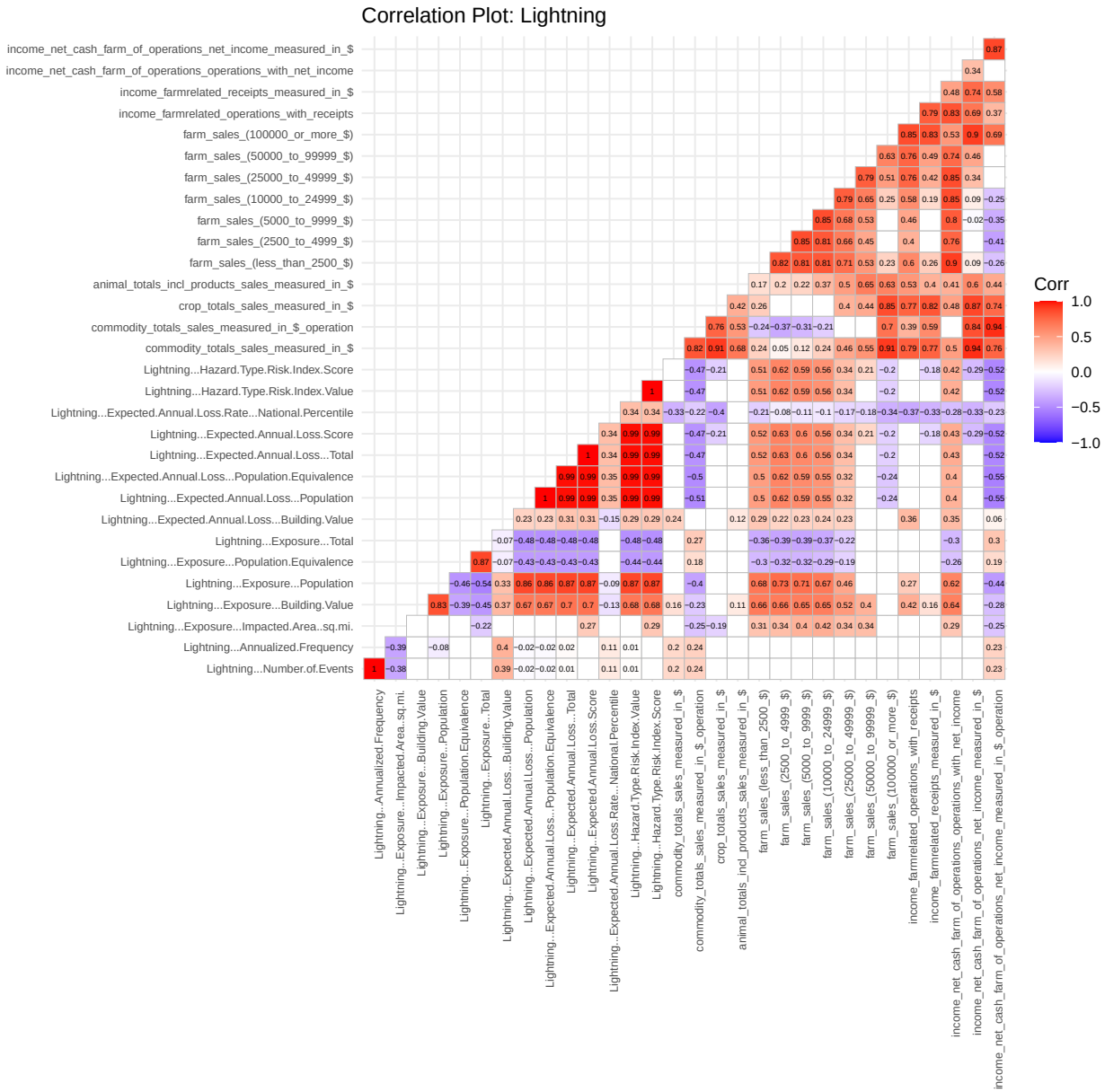

Correlation Plot: Ice Storm
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_farmrelated_receipts_measured_in_$
income_farmrelated_operations_with_receipts
farm_sales_(100000_or_more_$)
farm_sales_(50000_to_99999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(2500_to_4999_$)
farm_sales_(less_than_2500_$)
animal_totals_incl_products_sales_measured_in_$
crop_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
commodity_totals_sales_measured_in_$
Ice.Storm...Hazard.Type.Risk.Index.Score
Ice.Storm...Hazard.Type.Risk.Index.Value
Ice.Storm...Expected.Annual.Loss.Rate...National.Percentile
Ice.Storm...Expected.Annual.Loss.Rate...Building
Ice.Storm...Expected.Annual.Loss.Score
Ice.Storm...Expected.Annual.Loss...Total
Ice.Storm...Expected.Annual.Loss...Population.Equivalence
Ice.Storm...Expected.Annual.Loss...Population
Ice.Storm...Expected.Annual.Loss...Building.Value
Ice.Storm...Historic.Loss.Ratio...Buildings
Ice.Storm...Exposure...Total
Ice.Storm...Exposure...Population.Equivalence
Ice.Storm...Exposure...Population
Ice.Storm...Exposure...Building.Value
Ice.Storm...Exposure...Impacted.Area..sq.mi.
Ice.Storm...Annualized.Frequency
Ice.Storm...Number.of.Events
Corr
1.0
0.5
0.0
−0.5
−1.0

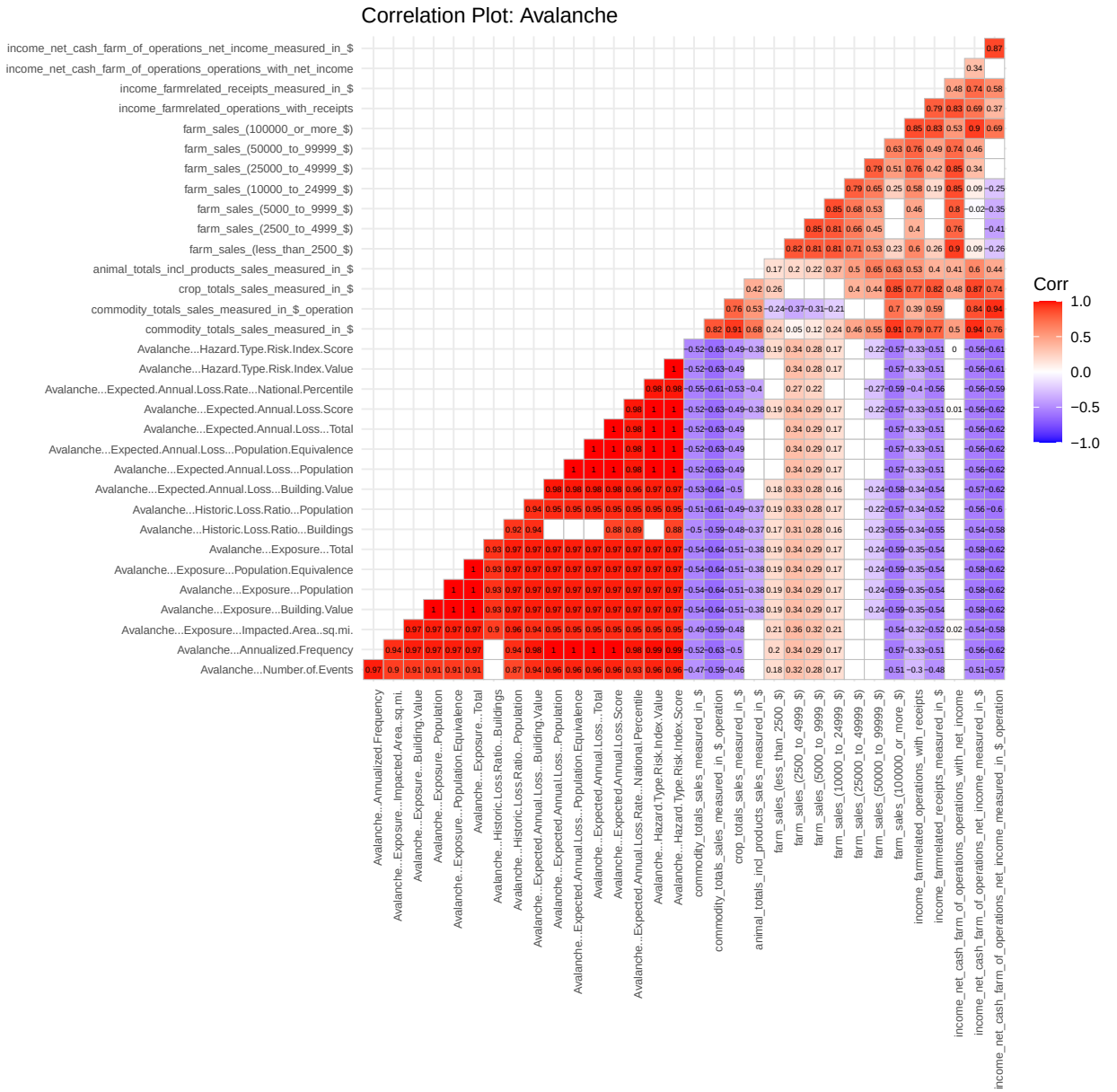
Correlation Plot: Strong Wind



Correlation Plot: Tornado




Correlation Plot: Lightning
Corr
1.0
0.5
0.0
−0.5
−1.0
Lightning...Number.of.Events
Lightning...Annualized.Frequency
Lightning...Exposure...Impacted.Area..sq.mi.
Lightning...Exposure...Building.Value
Lightning...Exposure...Population
Lightning...Exposure...Population.Equivalence
Lightning...Exposure...Total
Lightning...Expected.Annual.Loss...Building.Value
Lightning...Expected.Annual.Loss...Population
Lightning...Expected.Annual.Loss...Population.Equivalence
Lightning...Expected.Annual.Loss...Total
Lightning...Expected.Annual.Loss.Score
Lightning...Expected.Annual.Loss.Rate...National.Percentile
Lightning...Hazard.Type.Risk.Index.Value
Lightning...Hazard.Type.Risk.Index.Score
commodity_totals_sales_measured_in_$
commodity_totals_sales_measured_in_$_operation
crop_totals_sales_measured_in_$
animal_totals_incl_products_sales_measured_in_$
farm_sales_(less_than_2500_$)
farm_sales_(2500_to_4999_$)
farm_sales_(5000_to_9999_$)
farm_sales_(10000_to_24999_$)
farm_sales_(25000_to_49999_$)
farm_sales_(50000_to_99999_$)
farm_sales_(100000_or_more_$)
income_farmrelated_operations_with_receipts
income_farmrelated_receipts_measured_in_$
income_net_cash_farm_of_operations_operations_with_net_income
income_net_cash_farm_of_operations_net_income_measured_in_$
income_net_cash_farm_of_operations_net_income_measured_in_$_operation



Problems with Data Structure

Many of the columns are filled with NAs, but I am lead to believe this is because some factors simply do not happen in certain states/counties. This could also be due to how to data was transformed into a clean data set; further investigation is required. When processing the data prior to calculating correlations, I changed all NAs to 0, which may have caused faulty evaluation of correlations and/or significance as well. I chose to do this because many variables take the frequency or exposure from different variables, so if a county never gets any events of such then frequency, exposure, or any related factor is virtually nonexistent, and the data entry is entered as NA, and is therefore changed to zero by my data cleaning, which is why converting those value is a valid procedure. Changing NAs to zero also make it easier for R to calculate correlation coefficients. Keeping NA values can lead to the deletion of rows depending on the requirements of the algorithms used. Since each row is a different county, I want to avoid this as much as I can in order to have

the most complete data set possible (I have not removed any rows at this point). Mean/Median or K-Nearest Neighbor imputation was out of the question for this data. Doing so would have altered recorded fact and created falsehoods, adding bias that would otherwise not be present and negating the overall plausibility of the analysis. Removing rows without complete observations was not favorable for similar reasons. Since the purpose is to find correlations between variables and not related to machine learning at this time, the optimal method was converting NAs to 0 values to create complete observations.