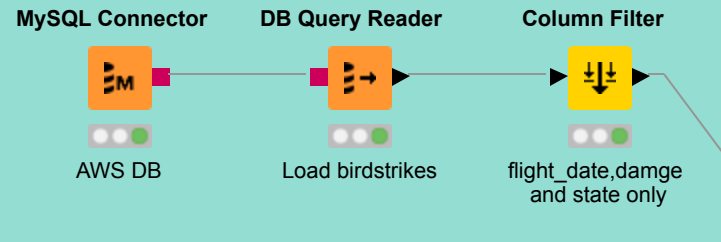In this workflow, we would like to study the correlation between the number of incidents in Birdstrikes DB and wind speed.
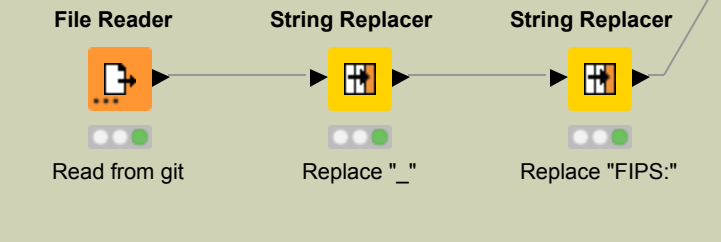
For this, we need to enrich every observation with average wind speed for the given date in the given state. The enrichment data source will be the Weather API of National Centers for Environmental Information.

This API requires a so called FIPS identifier which will be served as a CSV. This CSV contains FIPS identifiers for all US states with some errors …
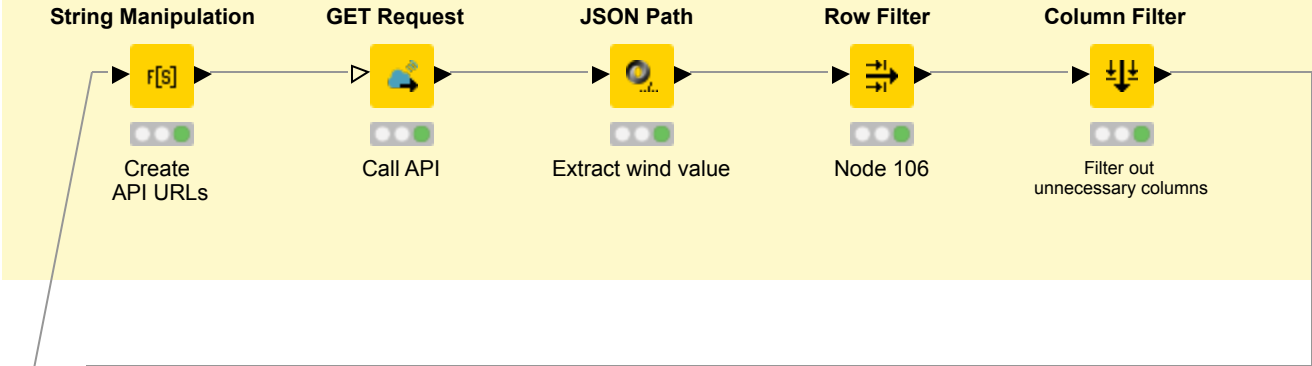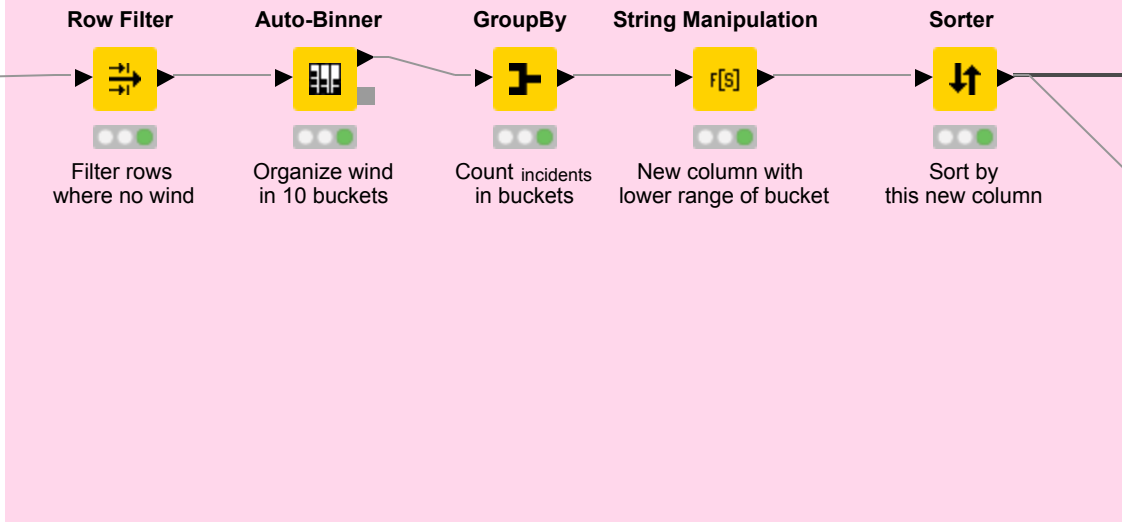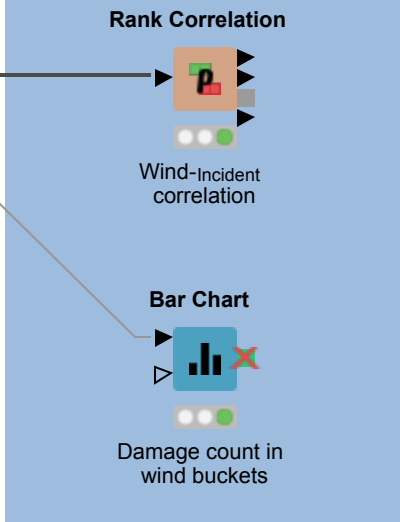
USE API TO ADD WIND SPEED

String Manipulation	GET Request	JSON Path	Row Filter	Column Filter
Create API URLs	Call API	Extract wind value	Node 106	Filter out unnecessary columns

LOAD BIRDSTRIKES

MySQL Connector	DB Query Reader	Column Filter
AWS DB	Load birdstrikes	flight_date,damge and state only

ADD STATE FIPS

File Reader	String Replacer	String Replacer
Read from git	Replace "_"	Replace "FIPS:"

Joiner

Adding fips to our datase

CLEAN AND TRANSFORM

Row Filter	Auto-Binner	GroupBy	String Manipulation	Sorter
Filter rows where no wind	Organize wind in 10 buckets	Count $_{incidents}$ in buckets	New column with lower range of bucket	Sort by this new column

ANALYTICS

Rank Correlation

Wind-$_{Incident}$ correlation

Bar Chart

Damage count in wind buckets

API URL Sample

https://www.ncdc.noaa.gov/cdo-web/api/v2/data?datasetid=GHCND&datatypeid=AWND&limit=1&startdate=2000-01-02&enddate=2000-01-02&locationid=FIPS:06