**Constraint Analyzer Project**

**Data Collection:**

* **Location of all the files mentioned in this document:**
* **S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data**

Needed data – Transmission outages data, Constraint data, Auction Contingency Data, Auction Mapping Document

\***For now this project is only developed for ERCOT region, once completed it can be developed for other regions as well.**

**\*For trial runs of the code data from the year 2018 would be used in the initial phase.**

**\*This document contains information about all the file extensions except .py extension for which documentation is done in CodeFiles\_Documentation.docx file**

**Source of Transmission Outage Data:**

* Location:
* S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint
* Data is taken from the years 2014 to 2019
* Yes Energy – data pulled using the Get\_Historical\_Trans\_Outage\_v1.py
* Link to .py file - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Get\_Historical\_Trans\_Outage\_v1.py

**\*Sample data folder contains the data files that will be used as trial data which would be used for testing the code file’s working. (Not used yet)**

**Preprocessing of Transmission Outages Files:**

Preprocessing of the files listed below must be done before using it as an input. All files are:

1. S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint\ERCOT\_TransmissionOutage\_2014-01-01\_to\_2014-12-31.xlsx
2. S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint\ERCOT\_TransmissionOutage\_2015-01-01\_to\_2015-12-31.xlsx
3. S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint\ERCOT\_TransmissionOutage\_2016-01-01\_to\_2016-12-31.xlsx
4. S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint\ERCOT\_TransmissionOutage\_2017-01-01\_to\_2017-12-31.xlsx
5. S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint\ERCOT\_TransmissionOutage\_2018-01-01\_to\_2018-12-31.xlsx
6. S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint\ERCOT\_TransmissionOutage\_2019-01-01\_to\_2019-07-24.xlsx

Following step are to be performed on all the files mentioned above:

1. Select ‘reported name’ column of the excel file.
2. Go to Data (Excel Ribbon) 🡪 Text to Column 🡪Select delimited 🡪 Click Next 🡪 Check Space (uncheck all others) 🡪 Click Next 🡪 Click on Destination (in Convert Text to Columns Wizard – Step 3 of 3) 🡪 Select 4 columns in the sheet after the last column i.e. ‘RELIABILITY\_SCORE’ 🡪 Click Finish
3. In the newly created columns, only keep column 1st and 3rd. Delete the other two columns.
4. Name the newly created 1st column as ‘from’ and 3rd column as ‘to’.
5. Also, copy this excel file - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents\2019.JUL.Monthly.Auction.MappingDocument.xlsx, sheet name – Lines and paste it to a new sheet of transmission outage file. (this must be done for all the transmission files)

**Source of Historical Constraint Data:**

* Location:
* S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Outages + Constraint
* Data is taken from the years 2015 to 2018.
* Energy Velocity is the vendor from where Historical Constraint data is pulled out.
* The reason for taking two different data sources for Constraint data and Transmission Outages data is that the mapping document used for mapping the Constraint-Contingency pair to the Transmission Lines in PowerWorld is pulled from ERCOT which maps more closely to Energy Velocity’s Constraint data than Yes Energy’s Constraint Data.
* Although, we are trying to pull constraint data from Yes energy as well and compare whether the above-mentioned fact is correct or not.
* Parameters considered in Historical Constraint Data:
* Monitored Element Name
* Contingency Name
* Local Date Time (Hour Beginning) with minutes
* Shadow Price ($/MWh)
* ISO Name
* Flowgate Name
* Time Zone
* Market Type
* Off/On Peak (entered manually by us)

All the columns other than the one mentioned above are deleted for all the years.

**Source of Mapping Document:**

* Location:
* S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents
* This mapping data is taken from ERCOT.
* The mapping document can be taken from any year (preferably latest year’s data should be taken) as it contains the list of constraint-contingency that have happened over the period.
* For now, we are using 2019 auction contingency document of July month so that we have data related to the latest constraint-contingency pair as well.
* Location of this data is: S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents

**Source/ Further Processing of uniquePairList.xlsx:**

* Location:
* S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\UniqueConstraintContingencyPair
* This file is created using a python file named ConstraintContingencyUniquePairCreation.py
* This file contains unique pairs of constraint-contingency for the year 2018.
* Sheet uniquePairList:
* Columns named – ISO Name, Flowgate, Monitored Element Name, Contingency Name, Local DateTIme (Hour Beginning) with Minutes, Time Zone, Market Type, Shadow Price $/MWh, OnPeak/Off Peak – EAST are fetched using the python code file.
* Columns named – Contingency Mapped From Bus Number, Contingency Mapped From Bus Name, Contingency Mapped To Bus Number, Contingency Mapped To Bus Name, Contingency Circuit ID, Monitored Line From Bus Number, Monitored Line From Bus Name, Monitored Line To Bus Number, Monitored Line To Bus Name, Monitored Circuit ID are fetched using VLOOKUP formula

For Contingency Data formula used - =IFERROR(VLOOKUP(D2,AuctionContingency2019JUL!$A:$I,5,FALSE)," ")

Where AuctionContingency2019JUL is Sheet 2 of uniquePairList.xlsx and the data of this sheet is taken from - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents\2019.JUL.Monthly.Auction.Contingencies.CSV

For Monitored Line Data formula used –

=IFERROR(VLOOKUP(C2,AuctionMapping2019JUL!$C:$J,4,FALSE)," ")

Where AuctionMapping2019JUL is Sheet 3 of uniquePairList.xlsx and the data of this sheet is taken from - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents\2019.JUL.Monthly.Auction.MappingDocument.xlsx

* Columns named

1. Interface Name is fetched using formula –

=IF(AND(J2<>" ",K2<>" ",L2<>" ",M2<>" ",O2<>"",P2<>" ",Q2<>" ",R2<>" "),CONCATENATE(P2,"-",R2," ","FLO"," ",K2,"-",M2),IF(AND(O2<>" ",P2<>" ",Q2<>" ",R2<>" "),CONCATENATE(P2,"-",R2," ","FLO"," ","BASE"),"NaN"))

1. Element a (Contingency) is fetched using formula-

=IF(AND(J2<>" ",K2<>" ",L2<>" ",M2<>" "),CONCATENATE("BRANCHOPEN"," ",J2," ",L2," ",O2),"NaN")

1. Element b (Monitored) is fetched using formula –

=IF(AND(O2<>" ",P2<>" ",Q2<>" ",R2<>" "),CONCATENATE("BRANCH"," ",O2," ",Q2," ",S2),"NaN")

1. Meter Far is given a value – “NO”
2. Weight is given a value – “1”

**Source of uniquePairList2015-2018.xlsx:**

* Location:
* S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\UniqueConstraintContingencyPair
* This file is created using a python file named ConstraintContingencyUniquePairCreation.py (second run)
* This file contains unique pairs of constraint-contingency for the years 2015-2018.
* Sheet uniquePairList:
* Columns named – ISO Name, Flowgate, Monitored Element Name, Contingency Name, Local DateTIme (Hour Beginning) with Minutes, Time Zone, Market Type, Shadow Price $/MWh, OnPeak/Off Peak – EAST are fetched using the python code file.
* Columns named – Contingency Mapped From Bus Number, Contingency Mapped From Bus Name, Contingency Mapped To Bus Number, Contingency Mapped To Bus Name, Contingency Circuit ID, Monitored Line From Bus Number, Monitored Line From Bus Name, Monitored Line To Bus Number, Monitored Line To Bus Name, Monitored Circuit ID are fetched using VLOOKUP formula

For Contingency Data formula used - =IFERROR(VLOOKUP(D2,AuctionContingency2019JUL!$A:$I,5,FALSE)," ")

Where AuctionContingency2019JUL is Sheet 2 of uniquePairList.xlsx and the data of this sheet is taken from - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents\2019.JUL.Monthly.Auction.Contingencies.CSV

For Monitored Line Data formula used –

=IFERROR(VLOOKUP(C2,AuctionMapping2019JUL!$C:$J,4,FALSE)," ")

Where AuctionMapping2019JUL is Sheet 3 of uniquePairList.xlsx and the data of this sheet is taken from - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents\2019.JUL.Monthly.Auction.MappingDocument.xlsx

* Columns named

1. Interface Name is fetched using formula –

=IF(AND(J2<>" ",K2<>" ",L2<>" ",M2<>" ",O2<>"",P2<>" ",Q2<>" ",R2<>" "),CONCATENATE(P2,"-",R2," ","FLO"," ",K2,"-",M2),IF(AND(O2<>" ",P2<>" ",Q2<>" ",R2<>" "),CONCATENATE(P2,"-",R2," ","FLO"," ","BASE"),"NaN"))

1. Element a (Contingency) is fetched using formula-

=IF(AND(J2<>" ",K2<>" ",L2<>" ",M2<>" "),CONCATENATE("BRANCHOPEN"," ",J2," ",L2," ",O2),"NaN")

1. Element b (Monitored) is fetched using formula –

=IF(AND(O2<>" ",P2<>" ",Q2<>" ",R2<>" "),CONCATENATE("BRANCH"," ",O2," ",Q2," ",S2),"NaN")

1. Meter Far is given a value – “NO”
2. Weight is given a value – “1”

**Source of PowerWorldFormat.xlsx:**

* Location - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\PowerWorldFormat.xlsx
* This file contains data only for the year 2018.
* Sheet name – Mapped
* This sheet contains the list of all the mapped contingency-monitored element pair i.e. for an interface name to be included either a monitored line from bus number was present or monitored line from bus number and contingency from bus number was present.
* All the columns of this sheet are created using code file named ConstraintContingencyMapping.py
* Sheet name – Unmapped
* This sheet contains list of all the unmapped contingency-monitored element pair i.e. a pair is included if a monitored line from bus number is not present and a contingency from bus number is present or a monitored line from bus number and contingency from bus number both are not present.
* All the columns in this sheet are created using a code file named ConstraintContingencyMapping.py

**Source of PowerWorldFormat2015-2018.xlsx:**

* Location - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\PowerWorldFormat.xlsx
* This file contains data only for the years 2015-2018.
* Sheet name – Mapped
* This sheet contains the list of all the mapped contingency-monitored element pair i.e. for an interface name to be included either a monitored line from bus number was present or monitored line from bus number and contingency from bus number was present.
* All the columns of this sheet are created using code file named ConstraintContingencyMapping.py
* Sheet name – Unmapped
* This sheet contains list of all the unmapped contingency-monitored element pair i.e. a pair is included if a monitored line from bus number is not present and a contingency from bus number is present or a monitored line from bus number and contingency from bus number both are not present.
* All the columns in this sheet are created using a code file named ConstraintContingencyMapping.py (second run).

**Source of all UniqueTransmissionOutagesMappedYEAR.xlsx (where YEAR can be – 2014,2015,2016,2017,2018,2019)**

* Location: S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data
* These files are created using a python file named – TransmissionOutagesMapping.py
* Rename sheet 1 as ‘Outages’
* Rename sheet 2 as AuctionMapping2019JUL\_LINES
* Copy - S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\Mapping Documents\2019.JUL.Monthly.Auction.MappingDocument.xlsx, sheet name = Autos and paste it into sheet 3 of this file, rename it as ‘AuctionMapping2019JUL\_AUTOS
* Delete the index column in ‘Outages’ sheet and ‘AuctionMapping2019JUL\_LINES sheet created by the python code.
* In the ‘Outages’ sheet Columns named – iso, facility, fromstation, tostation, kv, fromzone, tozone, facility\_type, type, type\_detail, status, status\_detail, startdate, enddate, planned\_startdate, planned\_enddate, open\_close, ticketid, facilityid, lastchangedate, publishdate, reported\_name, fromstationid, tostationid, reliability\_score, reliability\_accuracy, from, to are all fetched by the python code file itself.
* In the AuctionMapping2019JUL\_AUTOS sheet, insert new columns as follows:

1. Select ‘Operations\_Name’ column of the excel sheet.
2. Go to Data (Excel Ribbon) 🡪 Text to Column 🡪Select delimited 🡪 Click Next 🡪 Check Other – enter underscore symbol (uncheck all others) 🡪 Click Next 🡪 Click on Destination (in Convert Text to Columns Wizard – Step 3 of 3) 🡪 Select 2 columns in the sheet after the last column i.e. ‘ID’ 🡪 Click Finish
3. Name the 1st column as ‘Operations’ and delete the rest.
4. Insert a column next to ‘Operations\_Name’ column and move the newly created column ‘Operations’ next to it.
5. Insert a column next to ‘Operations’ named ‘Combine’ and use this formula =C2&"|"&E2

To obtain the values for this column.

* In the ‘Outages’ sheet Columns which are created manually after executing python code to obtain this file are:

1. combine: formula used for obtaining column value

=AA2&"|"&AB2

1. from\_bus\_number: formula used for obtaining column value =IF(H2="XFMR",IFERROR(VLOOKUP(AC2,AuctionMapping2019JUL\_AUTOS!$D:$J,3,FALSE)," "),IFERROR(VLOOKUP(AB2,AuctionMapping2019JUL\_LINES!$C:$J,4,FALSE)," "))
2. from\_bus\_name: formula used for obtaining column value =IF(H2="XFMR",IFERROR(VLOOKUP(AC2,AuctionMapping2019JUL\_AUTOS!$D:$J,4,FALSE)," "),IFERROR(VLOOKUP(AB2,AuctionMapping2019JUL\_LINES!$C:$J,5,FALSE)," "))
3. to\_bus\_number: formula used for obtaining column value =IF(H2="XFMR",IFERROR(VLOOKUP(AC2,AuctionMapping2019JUL\_AUTOS!$D:$J,5,FALSE)," "),IFERROR(VLOOKUP(AB2,AuctionMapping2019JUL\_LINES!$C:$J,6,FALSE)," "))
4. to\_bus\_name: formula used for obtaining column value =IF(H2="XFMR",IFERROR(VLOOKUP(AC2,AuctionMapping2019JUL\_AUTOS!$D:$J,6,FALSE)," "),IFERROR(VLOOKUP(AB2,AuctionMapping2019JUL\_LINES!$C:$J,7,FALSE)," "))
5. circuit\_id: formula used for obtaining column value =IF(H2="XFMR",IFERROR(VLOOKUP(AC2,AuctionMapping2019JUL\_AUTOS!$D:$J,7,FALSE)," "),IFERROR(VLOOKUP(AB2,AuctionMapping2019JUL\_LINES!$C:$J,8,FALSE)," "))
6. reported\_name\_duplicate: copy and paste column values of ‘reported\_name’ column.

* In the ‘AuctionMapping2019JUL\_LINES’ sheet create a new column next to the last column named ‘circuit\_id’ and name this new column as ‘outages\_reported\_names\_mapped’. The values for this column are obtained using the following formula:

=IFERROR(VLOOKUP(F2,Outages!$AD:$AI,6,FALSE)," ")

* In the ‘AuctionMapping2019JUL\_AUTOS’ sheet create a new column next to the last column named ‘circuit\_id’ and name this new column as ‘outages\_reported\_names\_mapped’. The values for this column are obtained using the following formula:

=IFERROR(VLOOKUP(F2,Outages!$AD:$AI,6,FALSE)," ")

**Source of UniqueTransmissionOutagesMapped2014-2019.xlsx:**

* Location: S:\asset ops\GO\_Group\Interns\2019\Anubha\Constraint Project\Data\UniqueTransmissionOutagesMapped2014-2019.xlsx
* This file is an output file which is generated by a python code file named – UniqueTransmissionOutagesConcat.py

**Why is mapping done?**

The data that we get from Yes Energy or Energy Velocity consists the names of Monitored Element and Contingency whereas PowerWorld only consists of numbers i.e. From Bus Number and To Bus Number. So, mapping of names to numbers needs to be done to perform calculations of sensitivity factors such as PTDF, LODF and TLR.

Similar process must be carried out for transmission outages to determine which transmission line is open/close in the power system network.

**Parameters that are thought to be considered in training dataset:**

* All transmission outages as different parameter (can have a 0/1 value or LODF)
* Load
* Fuel Price
* PTDF
* On Peak/Off Peak
* Generators
* Date
* Month
* Year
* Week
* Time
* Shadow Price