Proposal for Final Project

UCB Data Analytics Bootcamp

Hourly Electric Forecasts for Electric Grids in the U.S.

Prepared by:

Shandiz

Andrew DeVries

Dipesh Manandhar

Josh Bode

October 14, 2018

TABLE OF CONTENTS

TOPIC AND RATIONALE

The objective for this project is to take multiple years of historical hourly demand data for multiple electric system in the US, analyze the historical patterns, and develop predictive models to forecast the hourly demand for the next and week.

The electric market is a $350 billion business per year in the U.S. alone. The hourly forecasts are critical because they drive infrastructure investments and market prices spike when demand is high and available capacity is low. Currently, each electric system (aka balancing authority) develop its own forecast. Both the historical actual and forecast hourly data is publicly available via and API.

Electric demand is highly sensitive to weather, hour of day (occupancy schedule), day of week, and season. The weather varies by location but so does the weather sensitivity. Weather sensitivity is driven by the penetration of air conditioning, electric heating, and the mix of industrial/commercial facilities. As a result, the model will be different for each of the locations.

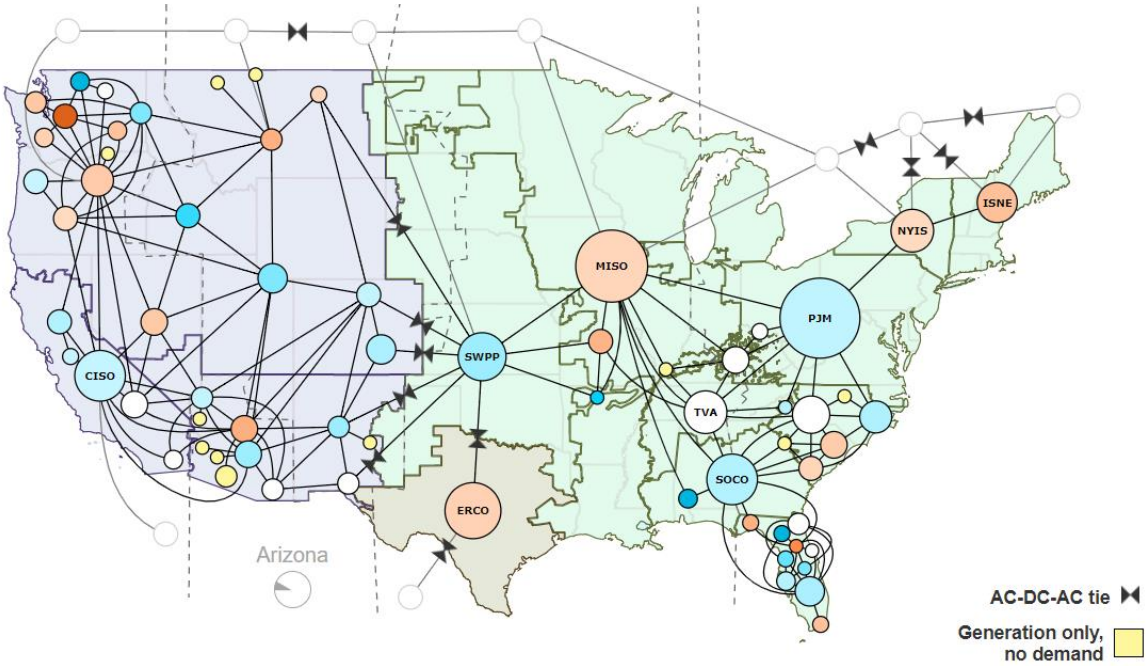
DATASETS

Data	Description	Link
Map of electric systems	Lat/long for all the major US electric grids	https://www.eia.gov/realtime_grid/#/status?end=20181013T13
Historical actual data by operating system	▪ System ▪ Date ▪ Hour ▪ MW (demand)	https://www.eia.gov/opendata/qb.php?category=2122628
Historical forecast data by operating system	▪ System ▪ Date ▪ Hour ▪ MW (demand)	https://www.eia.gov/opendata/qb.php?category=2122627
Weather station locations and characteristics	▪ Name and IDs ▪ Location, including lat/lon ▪ Dates for which weather station data is available	https://www1.ncdc.noaa.gov/pub/data/noaa/isd-history.csv
Historical weather data	▪ Station ▪ Date and hour ▪ Temperature (dry and wet bulb) ▪ Humidity ▪ Precipitation (rain/snow)	https://www1.ncdc.noaa.gov/pub/data

	▪ Wind direction and speed
Forecast weather data	▪ Nearest station ▪ Date and hour ▪ Forecasted hourly temperature (7 days of forecast data)

SCREEN SHOTS OF THE DATA

MAP OF ELECTRIC SYSTEMS IN US



HISTORICAL ACTUAL DATA

API Query Browser

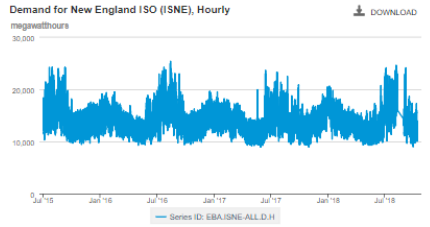
EIA Data Sets > U.S. Electric System Operating Data > Demand

API CALL TO USE: http://api.eia.gov/series/?api_key=YOUR_API_KEY_HERE&series_id=EBA.ISNE-ALL.D.H

SERIES NAME: Demand for New England ISO (ISNE), Hourly

SERIES ID: EBA.ISNE-ALL.D.H
Show me how to embed a chart of this series

GEOSET ID: EBA.D.H
Show me how to embed a map of this set



Source: U.S. Energy Information Administration

CHART DATA

Series Name	Period	Frequency	Value	Units
Demand for New England ISO (ISNE), Hourly	20181013T16Z	H	12099	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T17Z	H	12479	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T16Z	H	12769	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T15Z	H	12769	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T14Z	H	12471	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T13Z	H	11807	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T12Z	H	11227	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T11Z	H	10926	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T10Z	H	10346	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T09Z	H	10153	megawatthours
Demand for New England ISO (ISNE), Hourly	20181013T08Z	H	10099	megawatthours

HISTORICAL DAY AHEAD FORECAST DATA (OUR COMPETITION)

API Query Browser

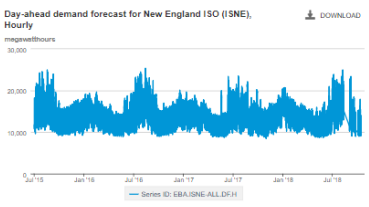
EIA Data Sets > U.S. Electric System Operating Data > Day-ahead demand forecast

API CALL TO USE: http://api.eia.gov/series/?api_key=YOUR_API_KEY_HERE&series_id=EBA.ISNE-ALL.DF.H

SERIES NAME: Day-ahead demand forecast for New England ISO (ISNE), Hourly

SERIES ID: EBA.ISNE-ALL.DF.H
Show me how to embed a chart of this series

GEOSET ID: EBA.DF.H
Show me how to embed a map of this set



Source: U.S. Energy Information Administration

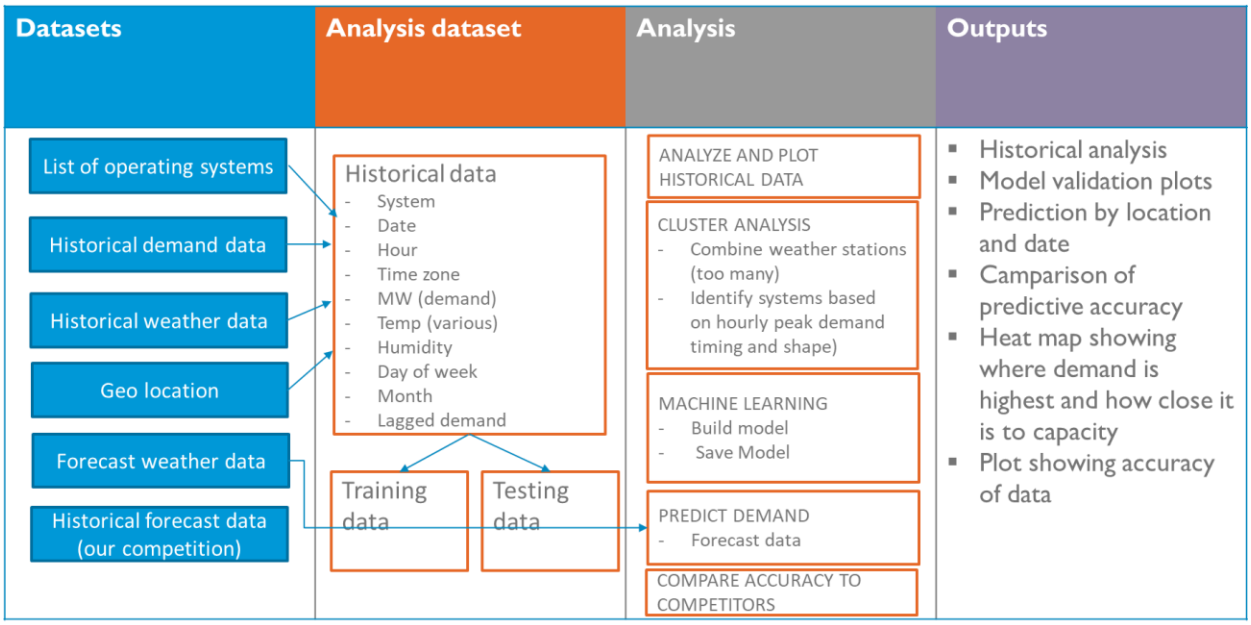
CHART DATA

Series Name	Period	Frequency	Value	Units
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181014T04Z	H	10530	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181014T03Z	H	11330	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181014T02Z	H	12130	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181014T01Z	H	12770	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181014T00Z	H	13380	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181013T23Z	H	13500	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181013T22Z	H	12720	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181013T21Z	H	12160	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181013T20Z	H	11730	megawatthours
Day-ahead demand forecast for New England ISO (ISNE), Hourly	20181013T19Z	H	11700	megawatthours

WEATHER STATION CHARACTERISTICS AND LOCATIONS

	A	B	C	D	E	F	G	H	I	J	K
1	USAF	WBAN	STATION N	CTRY	STATE	ICAO	LAT	LON	ELEV(M)	BEGIN	END
2	7018	99999	WXPOD 7018				0	0	7018	20110309	20130730
3	7026	99999	WXPOD 70	AF			0	0	7026	20120713	20170822
4	7070	99999	WXPOD 70	AF			0	0	7070	20140923	20150926
5	8260	99999	WXPOD8270				0	0	0	20090101	20100731
6	8268	99999	WXPOD82	AF			32.95	65.567	1156.7	20100519	20120323
7	8307	99999	WXPOD 83	AF			0	0	8318	20100421	20100421
8	8411	99999	XM20							20160217	20160217
9	8414	99999	XM18							20160216	20160217
10	8415	99999	XM21							20160217	20160217
11	8418	99999	XM24							20160217	20160217
12	10000	99999	BOGUS NO	NO		ENRS				20010927	20051231
13	10010	99999	JAN MAYE	NO		ENJA	70.933	-8.667	9	19310101	20181009
14	10013	99999	ROST	NO						19861120	19880105
15	10014	99999	SORSTOKK	NO		ENSO	59.792	5.341	48.8	19861120	20181009
16	10015	99999	BRINGELA	NO			61.383	5.867	327	19870117	20081231
17	10016	99999	RORVIK/R	NO			64.85	11.233	14	19870116	19910806
18	10017	99999	FRIGG	NO		ENFR	59.98	2.25	48	19880320	20050228
19	10020	99999	VERLEGEN	NO			80.05	16.25	8	19861109	20181009
20	10030	99999	HORNSUN	NO			77	15.5	12	19850601	20181009
21	10040	99999	NY-ALESU	NO		ENAS	78.917	11.933	8	19730101	20140523
22	10050	99999	ISFJORD R	SV			78.067	13.633	9	19310103	20051109
23	10060	99999	EDGEOYA	NO			78.25	22.817	14	19730101	20181009
24	10070	99999	NY-ALESU	SV			78.917	11.933	7.7	19730106	20181009
25	10071	99999	LONGYEAF	SV			78.217	15.583	37	20050210	20050210
26	10080	99999	LONGYEAF	SV		ENSB	78.246	15.466	26.8	19750929	20181009
27	10090	99999	KARL XII O	SV			80.65	25	5	19550101	20181009
28	10100	99999	ANDOYA	NO		ENAN	69.293	16.144	13.1	19310103	20181009
29	10110	99999	KVITOYA	SV			80.067	31.5	10	19861118	20181009

PROJECT WORKFLOW



SCALING STRATEGY

- Phase 1 – Build out one location using static data
- Phase 2 – Run analysis for multiple systems
- Phase 3 – Automate pulling of updated historical data, forecast weather, and models

- Phase 4 – Convert into an app that automatically updates periodically

Initial visualization will be in Python and Tableau. If time permits, build out site.

TEAM MEMBER ROLES

Task	Key Steps	Team Member
Data engineering	▪ Download all historical demand ▪ Download historical weather data – 100 station in US for 5 years (in locations) ▪ Download the historical forecast demand data ▪ Standardize the time zone and daylight saving time issues ▪ Produce feature for different weather metrics, day of week, etc.	▪ Dipesh ▪ Shandiz ▪ Andrew ▪ Dipesh/Shandiz/Andrew ▪ Josh
Data Analysis	▪ Historical data analysis/visualization ▪ Cluster analysis for weather ▪ Cluster analysis of peak day shape ▪ Machine learning model ▪ Prediction using machine learning model ▪ Accuracy metrics for our models versus competitor models	▪ Andrew ▪ Shandiz ▪ Dispesh ▪ Josh/Dispesh/Shandiz/Andrew ▪ Josh/Dispesh/Shandiz/Andrew ▪ Josh
Outputs	▪ Historical analysis ▪ Model validation plots (predicted v actual) ▪ Predicted demand by date hour and location ▪ Maps showing where demand is highest and close it is to capacity ▪ Plots comparing the accuracy our predictions to the competition	▪ Andrew ▪ Josh ▪ Team effort ▪ Andrew/Josh ▪ Josh/Dispesh
Scaling and Automation	▪ Fully automate all aspects of data pulls ▪ Schedule data extraction using the API's	

- Schedule update to the machine learning models
- Deploy