Software and
Programming in IoT

ENERGY CHANGE DETECTION

Table of CONTENTS

Project OVERVIEW

- Dataset used: smart-home-dataset-with-weather-information
- In this dataset, energy from the house appliances and weather information are provided.
- A few columns of this dataset are: Amount of house energy consumption, Dishwasher, Furnace, Home office, Fridge etc.
- Data is collected using a Smart Meter
- By using this dataset, we can understand the relationship between energy consumption by appliances and weather changes while detecting anomalous usage of appliances.

DATA COLLECTION AND PREPROCESSING

1 Data column cleaning

2 Missing value check

3 Data info and Potential Cleaning

4 Time conversion and feature extraction

5 Day part classification

6 Correalation analysis

DATA COLUMN CLEANING

- THE INITIAL CELL EXAMINES THE COLUMNS AND REMOVES THE TEXT '[KW]' FROM THE COLUMN NAMES FOR CLEANER LABELS
- IT CREATES TWO COLUMNS BY SUMMING SPECIFIED SUB-COLUMNS
- THEN, IT DROPS THE NOW-REDUNDANT INDIVIDUAL COLUMNS, AS WELL AS UNNEEDED COLUMNS

COLUMN NAMES:

```
df['Furnace'] = df[['Furnace 1','Furnace 2']].sum(axis=1)
df['Kitchen'] = df[['Kitchen 12','Kitchen 14','Kitchen 38'
df.drop(['Furnace 1','Furnace 2','Kitchen 12','Kitchen 14'
```

```
(['time', 'use [kW]', 'gen [kW]', 'House overall [kW]', 'Dishwasher [kW]',
  'Furnace 1 [kW]', 'Furnace 2 [kW]', 'Home office [kW]', 'Fridge [kW]',
  'Wine cellar [kW]', 'Garage door [kW]', 'Kitchen 12 [kW]',
  'Kitchen 14 [kW]', 'Kitchen 38 [kW]', 'Barn [kW]', 'Well [kW]',
  'Microwave [kW]', 'Living room [kW]', 'Solar [kW]', 'temperature',
  'icon', 'humidity', 'visibility', 'summary', 'apparentTemperature',
  'pressure', 'windSpeed', 'cloudCover', 'windBearing', 'precipIntensity',
  'dewPoint', 'precipProbability'],
 dtype='object')
```

MISSING VALUE CHECK

- THE CODE IDENTIFIES MISSING VALUES AND REMOVES THE LAST ROW TO POTENTIALLY ADDRESS INDEXING OR DATA QUALITY ISSUES

ROW CONTAINING NAN VALUES:

	time	use	gen	House overall	Dishwasher	Home office	Fridge	Wine cellar	Garage door	Barn	...	apparentTemperature	pressure	windSpeed	cloudCover	windBearing	precipIntensity	dewPoint
503910	\	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN

rows × 27 columns

DATA INFO AND POTENTIAL CLEANING

- DF.INFO() PROVIDES AN OVERVIEW OF DATA TYPES AND MISSING VALUES
- IT EXPLORES THE UNIQUE VALUES IN THE CLOUDCOVER COLUMN TO CHECK FOR ANOMALIES, WITH A SPECIFIC CHECK FOR OCCOURENCES OF A PLACEHOLDER 'CLOUDCOVER'
- THE ANOMALOUS VALUES ARE REPLACED WITH THE LAST VALID VALUE AND THE COLUMN IS CONVERTED INTO A FLOAT

```
20  cloudCover            503910 non-null  object
21  windBearing           503910 non-null  float64
22  precipIntensity       503910 non-null  float64
```

```
df['cloudCover'].replace(['cloudCover'], method='bfill', inplace=True)
df['cloudCover'] = df['cloudCover'].astype('float')
```

TIME CONVERSION AND FEATURE EXTRACTION

- CONVERTS THE TIME COLUMN TO A DATETIME FORMAT
- EXTRACTS SEVERAL DATE-RELATED FEATURES TO CREATE TIME-BASED BREAKDOWN IN SUBSEQUENT ANALYSIS

use	gen	House overall	Dishwasher	Home office	Fridge	Wine cellar	Garage door	Barn	...	precipProbability	Furnace	Kitchen	year	month	day	weekday	weekofyear	hour	minute
833	0.003483	0.932833	0.000033	0.442633	0.124150	0.006983	0.013083	0.031350	...	0.0	0.082617	0.000567	2016	1	1	Friday	53	5	0
333	0.003467	0.934333	0.000000	0.444067	0.124000	0.006983	0.013117	0.031500	...	0.0	0.084533	0.000567	2016	1	1	Friday	53	5	1
817	0.003467	0.931817	0.000017	0.446067	0.123533	0.006983	0.013083	0.031517	...	0.0	0.083017	0.000617	2016	1	1	Friday	53	5	2

DAY PART CLASSIFICATION

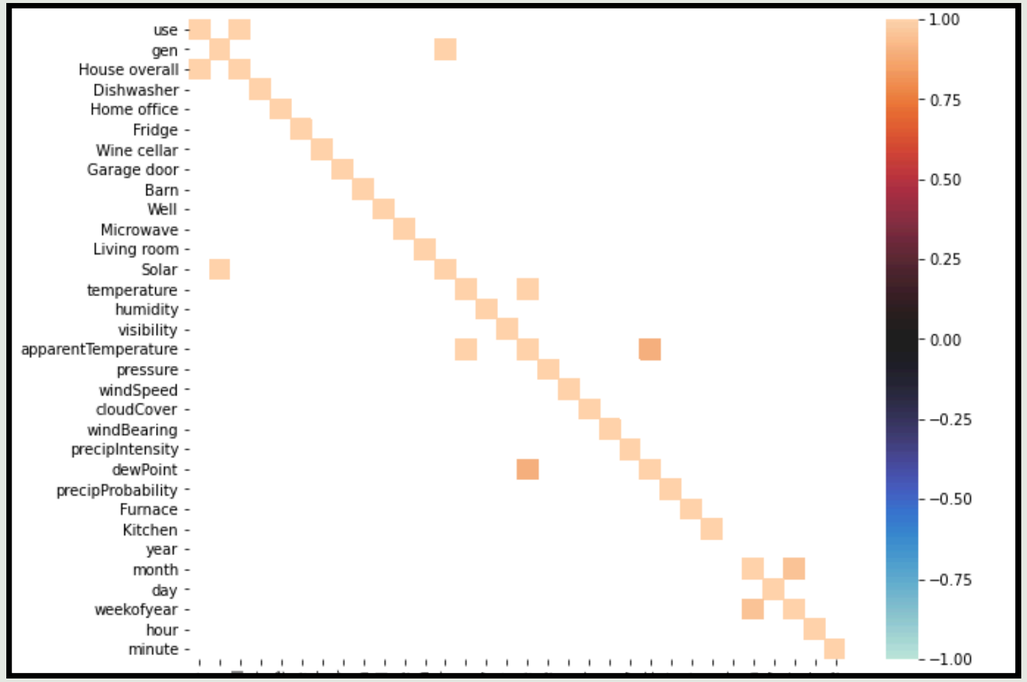
- DEFINES A FUNCTION THAT CATEGORIZES EACH HOUR INTO A DAY PART BASED ON TYPICAL ENERGY USAGE PATTERNS
- APPLIES THIS FUNCTION TO CREATE A NEW COLUMN IN THE DATAFRAME FOR USING IN ANALYZING USAGE BY TIME OF DAY

year	month	day	weekday	weekofyear	hour	minute	timing
2016	1	1	Friday	53	5	0	Morning
2016	1	1	Friday	53	5	1	Morning
2016	1	1	Friday	53	5	2	Morning

```
def hours2timing(x):
    if x in [22,23,0,1,2,3]:
        timing = 'Night'
    elif x in range(4, 12):
        timing = 'Morning'
    elif x in range(12, 17):
        timing = 'Afternoon'
    elif x in range(17, 22):
        timing = 'Evening'
    else:
        timing = 'X'
    return timing
```

CORRELATION ANALYSIS

- IT GENERATES A HEATMAP OF CORRELATIONS BETWEEN COLUMNS WITH SIGNIFICANT RELATIONSHIPS
- THIS STAT HELPS IDENTIFY CLOSELY RELATED FEATURES, PROVIDING INSIGHTS INTO HOW DIFFERENT ENERGY USAGES CORRELATE IN THE SMART HOME SETTING



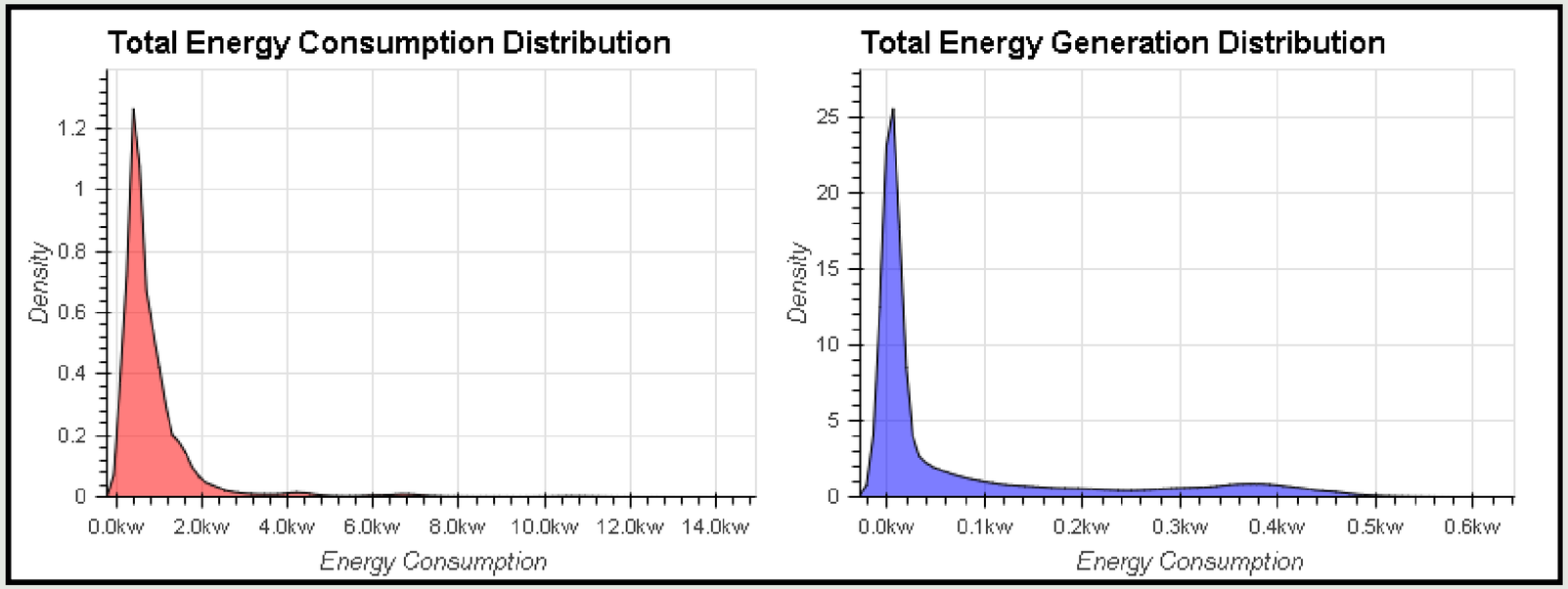

```
df['use_HO'] = df['use']
df['gen_Sol'] = df['gen']
df.drop(['use','House overall','gen','Solar'], axis=1,
df.head(3)
```

EXPLORATORY DATA ANALYSIS (EDA)

1. Distributions
2. Aggregations
3. Time-Series Analysis
4. Correlation Analysis
5. Goal

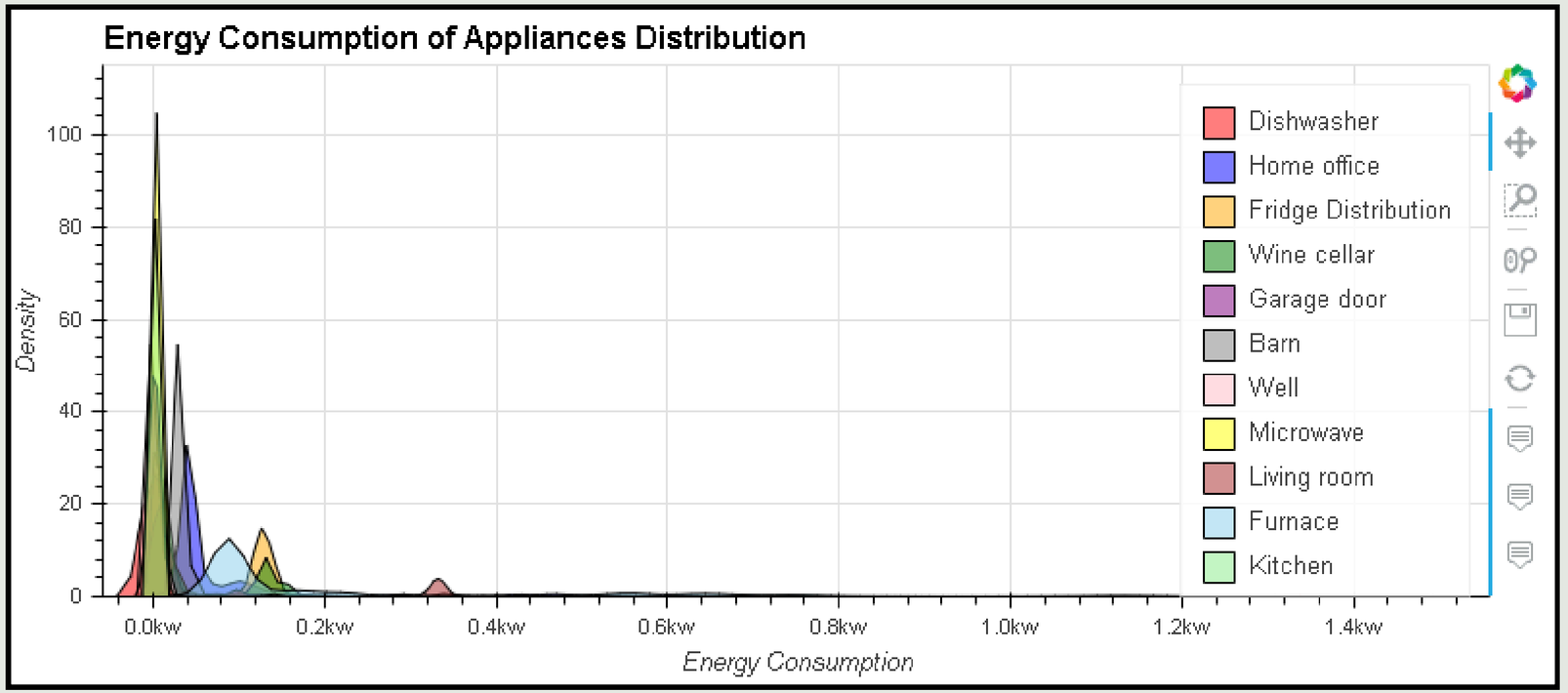
ENERGY DISTRIBUTIONS

- **DISTRIBUTIONS OF TOTAL ENERGY CONSUMPTION AND GENERATION AS WELL AS INDIVIDUAL APPLIANCES ARE PLOTTED USING HOLOVIEWS TO SHOW DENSITY DISTRIBUTIONS BY APPLIANCE AND ENERGY USAGE, AIDING IN UNDERSTANDING THE SPREAD AND PEAK USAGE TIMES FOR VARIOUS DEVICES.**



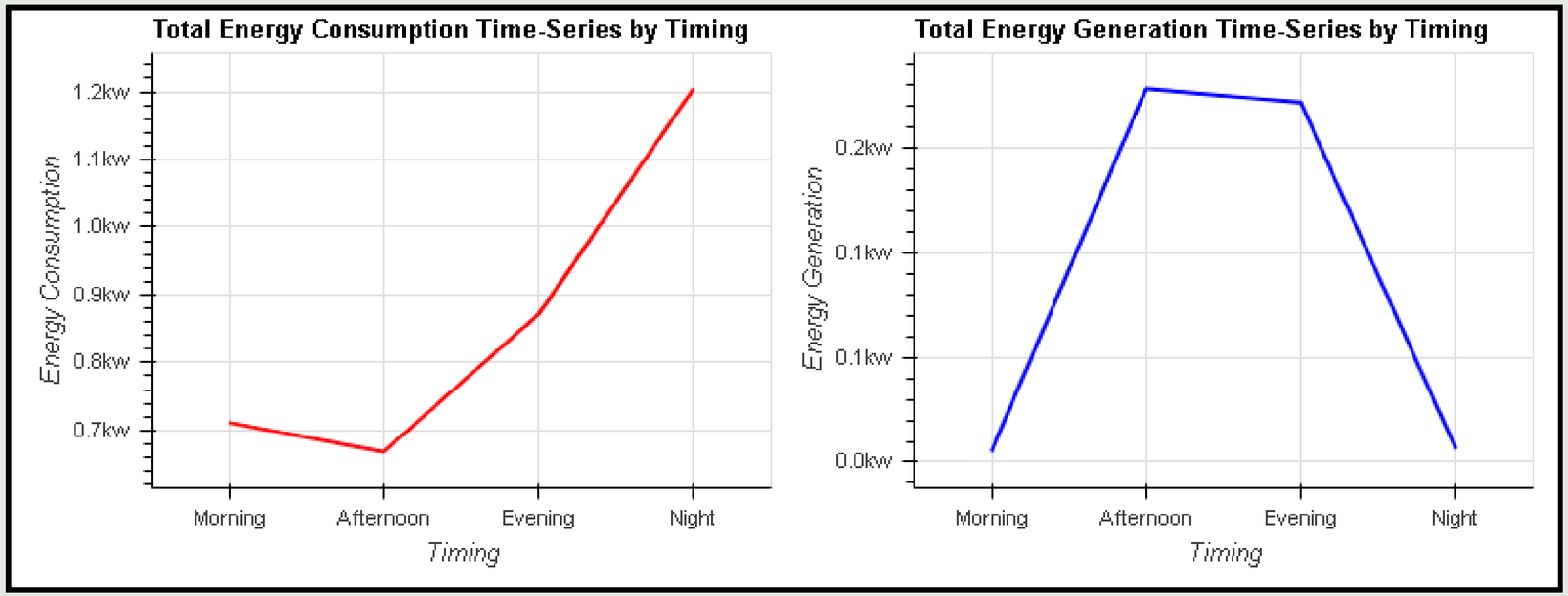
AGGREGATIONS

- FUNCTIONS ARE DEFINED TO AGGREGATE ENERGY CONSUMPTION AND GENERATION DATA BY MONTH, WEEKDAY, AND TIME-OF-DAY SEGMENTS. THE GOAL HERE IS TO OBSERVE CONSUMPTION AND GENERATION TRENDS OVER TIME, WITH VISUALIZATIONS FOR EACH TIME SEGMENT AND OVERALL APPLIANCE USAGE.



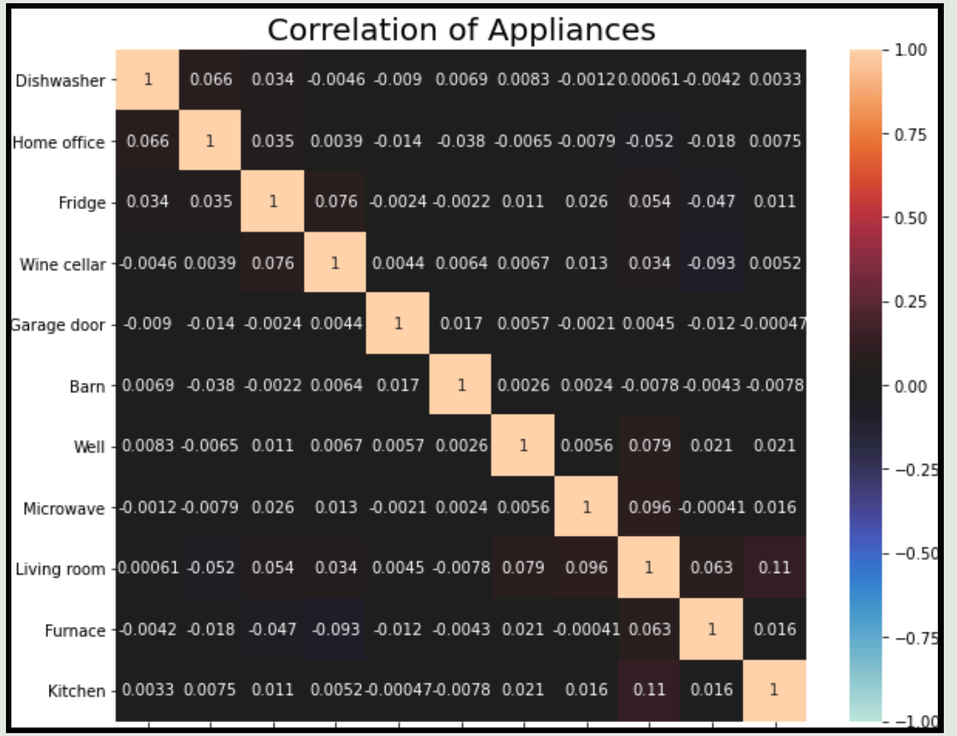
TIME-SERIES ANALYSIS

- DAILY AND MONTHLY TIME-SERIES CURVES ARE GENERATED FOR ENERGY CONSUMPTION, ENERGY GENERATION, AND SPECIFIC APPLIANCES, OFFERING INSIGHTS INTO HOW USAGE VARIES ACROSS THESE TIMESCALES. TEMPERATURE AND HUMIDITY TIME-SERIES ARE ALSO PLOTTED TO CAPTURE SEASONAL TRENDS AND WEATHER CONDITIONS OVER TIME.



CORRELATION ANALYSIS

- USING HEATMAPS, THE CORRELATION BETWEEN APPLIANCE ENERGY USAGE IS EXAMINED TO IDENTIFY RELATIONSHIPS AND DEPENDENCIES, WHILE ANOTHER HEATMAP EXPLORES WEATHER VARIABLE CORRELATIONS. A COMPREHENSIVE HEATMAP INCLUDES BOTH APPLIANCES AND WEATHER VARIABLES, AIMING TO FIND POTENTIAL RELATIONSHIPS BETWEEN ENVIRONMENTAL FACTORS AND HOUSEHOLD ENERGY PATTERNS.



GOAL

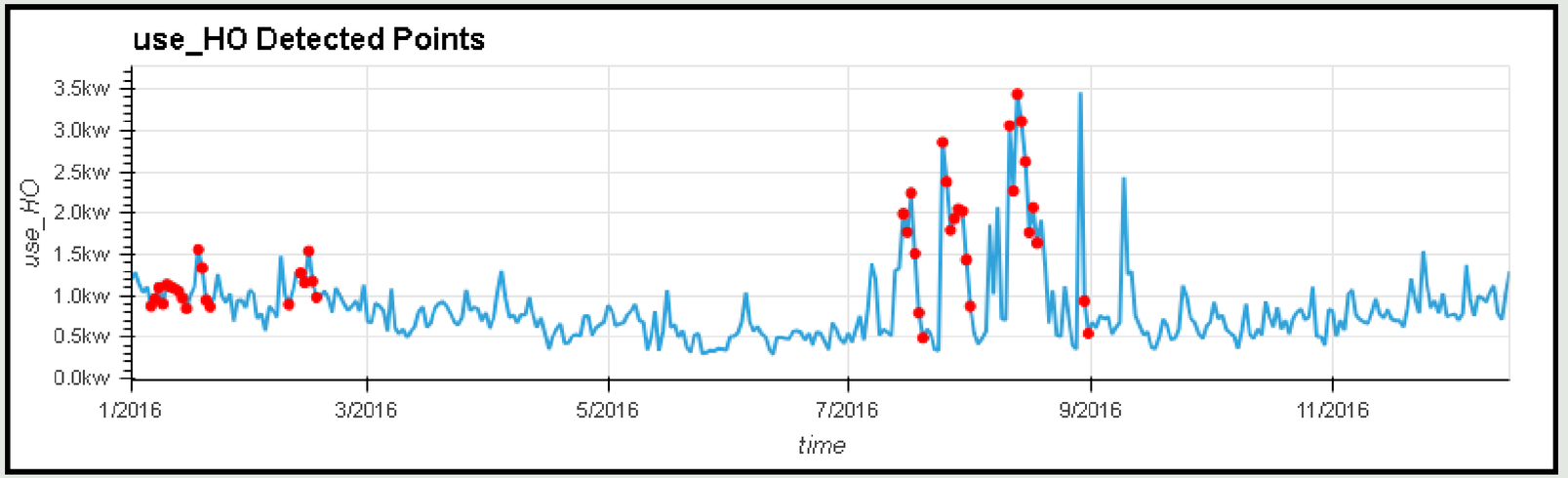
- THIS ANALYSIS SEEKS TO FIND PATTERNS IN ENERGY USAGE AND POTENTIAL CORRELATIONS WITH WEATHER FACTORS, HELPING TO INFORM PREDICTIVE MAINTENANCE, ENERGY-SAVING MEASURES, OR AUTOMATION STRATEGIES BASED ON ENVIRONMENTAL AND USAGE DATA TRENDS.

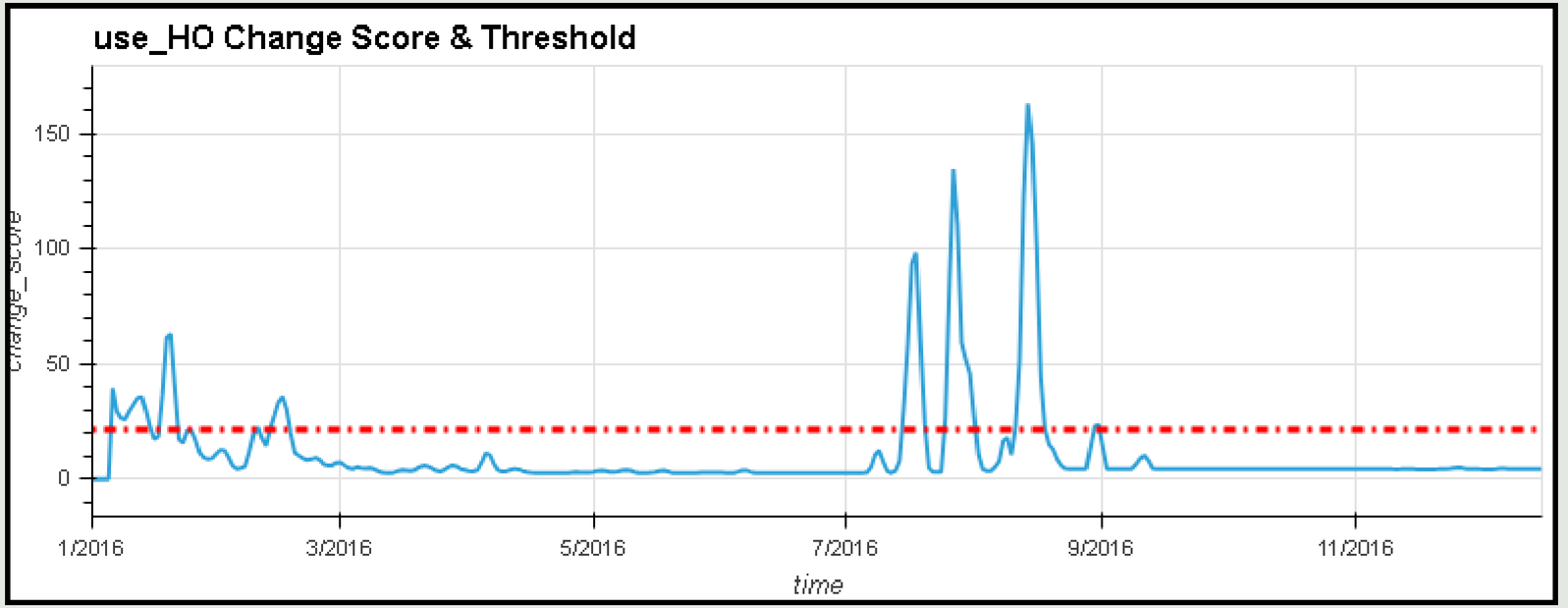
```
fig,ax = plt.subplots(figsize=(20, 12))
corr = df[['use_HO','gen_Sol','Dishwasher','Home office','Fridge','Wine cellar','Garage door','Barn','Well','Microwave','Living room','Furnace'
           'temperature','apparentTemperature','humidity','visibility','pressure','windSpeed','cloudCover','precipIntensity','dewPoint']].corr
sns.heatmap(corr, annot=True, vmin=-1.0, vmax=1.0, center=0)
ax.set_title('Correlation of Appliances & Weather Information',size=20)
plt.show()
```

MODELING

THE CHANGE DETECTION FUNCTION

- THE `CHNG_DETECTION` FUNCTION DETECTS CHANGE POINTS IN A TIME SERIES USING THE CHANGEFINDER ALGORITHM
- IT RESAMPLES THE DATA, COMPUTES A CHANGE SCORE BASED ON THE SEQUENTIALLY DISCOUNTING AR MODEL, AND SMOOTHS THESE SCORES.
- THE FUNCTION THEN IDENTIFIES SIGNIFICANT CHANGES BY CALCULATING AN UPPER THRESHOLD AND VISUALIZES BOTH THE CHANGE SCORES AND THE ORIGINAL DATA, HIGHLIGHTING THE DETECTED CHANGE POINTS.
- THIS HELPS IN DISTINGUISHING BETWEEN TEMPORARY ANOMALIES AND SUSTAINED TREND SHIFTS, PROVIDING INSIGHTS INTO UNDERLYING CHANGES IN THE DATA, SUCH AS CONSUMPTION PATTERNS.


use_HO Detected Points
3.5kw
3.0kw
2.5kw
2.0kw
1.5kw
1.0kw
0.5kw
0.0kw
use_HO
1/2016
3/2016
5/2016
7/2016
9/2016
11/2016
time


use_HO Change Score & Threshold
150
100
50
0
Change_Score
1/2016
3/2016
5/2016
7/2016
9/2016
11/2016
time

KEY USE CASES OF CHANGE FINDER

- ANOMALY DETECTION: FINDING ABNORMAL PATTERNS IN METRICS LIKE ENERGY USAGE, FINANCIAL TRANSACTIONS, OR WEB TRAFFIC.
- TREND CHANGE DETECTION: DETECTING SHIFTS IN TREND BEHAVIOR FOR PREDICTIVE MAINTENANCE OR ECONOMIC ANALYSIS.
- EVENT DETECTION: IDENTIFYING SPECIFIC EVENTS IN TIME-SERIES DATA, LIKE AN INCREASE IN TEMPERATURE OR SYSTEM FAILURES.

FUTURE SCOPE OF PROJECT

- FORECASTING FOR PRIOR ALERTS, (EG. GOOGLE HOME VOICE BASED RELAYING)
- ALERT SCHEMES, (EG. USING GSM OR GPRS RELATED MODULES TO RELAY MESSAGES)
- HARDWARE RELATED SOLUTIONS, (EG. ENERGY CONSUMPTION CONTROL AND PREDICTIVE MAINTENANCE)

    - EXAMPLE USE CASE:
    - SMART HOME: A SMART METER WAS USED TO COLLECT THE DATA, THE SAME CAN BE IMPLEMENTED TO MONITOR AND DETECT CHANGES.
    - THIS CAN HELP IMPLEMENT TARGETED CONTROL OF ENERGY CONSUMPTION
    - SEND IN ALERTS THROUGH SMS AND RELAY NECESSARY INFORMATION (APPLIANCE DESCRIPTION)

REFERENCES

- DATA UNDERSTANDING & EDA & TIME-SERIES MODELING
- CHANGEFINDER LIBRARY USAGE
- CHANGE DETECTION USING SDAR ALGORITHM PAPER