Final Presentation:
UBCO MDS Capstone - Urban Data Lab (UDL)

June 23, 2020

Connor Lee, Claudia Nikel, Eva Nguyen, Alex Tamm

Outline

- Project Background
- Overview of the Project
  - Research Question
  - Proposed Solutions & Deliverables
- The Data
  - What the data looks like & why it's important
  - High level overview
- Project Approach
  - Problem Definition
  - Our approach
- Project Details
- Results & Analysis
- Conclusion
  - Lessons Learned
  - Difficulties
- Acknowledgments
- Questions

Project Background

Urban Data Lab (UDL) Capstone Client Overview

Contacts are:

- Jiachen Wei (MDS alum)
- Mike Kennedy, Ph.D

Client Overview

- Formed in 2019 to advance data analytics capabilities and open data access on UBC Vancouver campus to address campus-wide sustainability challenges

Data Overview

- UDL mirrored and stored live-streaming building energy datasets in InfluxDB and made it available to students and researchers
- The SKYSPARK database provides data recorded by the meters and smart devices of many UBC buildings

Overview of Project

Overview

Research Question

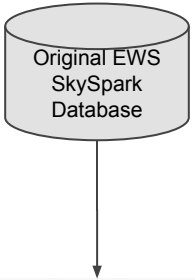
- Based on a building's sensor data, how can the data be grouped automatically into end-use classifications?

Proposed Solution and Deliverables

- A Python program that:
  - Queries and cleans the data required for classifying instrumentation by end-use for the Pharmacy building
  - Classifies instruments by end-use
  - Web-based Grafana dashboard
- Final report & presentation to UDL

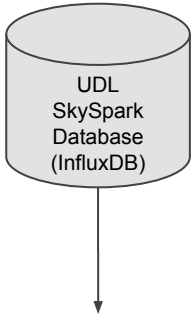
The Data

What does the data look like?



id	ahu	ahuMode	air	alarm	avg	bacnetConnRef	bacnetCur	bacnetHis	bacnetObjectId	bac
Pharmacy Heating Plant HX-2 P-HX2A HX2_PHX2A_VFD_PWR(kW...						PHARMA PHARM_HX2_FCU_B05 (701100)		TL56		✓
Pharmacy Elec Submeters LEED-2N1PC3 2N1PC3_CurrentC						PHARMA PHARM_LEED_METER_MOD2 (702100)		TL59		✓
Pharmacy Rm Corr FC-513 FCU_513_S			✓			PHARMA PHARM_FCU_513 (700742)		TL1		✓
Pharmacy Rm 1420 FC-111 FCU_111_S			✓			PHARMA PHARM_FCU_111/FF_102/EF_106 (701206)		TL1		✓
Pharmacy Elec Submeters LEED-6ETLE1 6ETLE1_CurrentB						PHARMA PHARM_LEED_METER_MOD1 (702000)		TL208		✓
Pharmacy Rm B503 EAV-BS048 BS048_AVG_SPACE_TEMP_AV			✓			PHARMA PHARM_MACRO_SERVER_FH_TEMP_RM (702900)		TL182		✓
Pharmacy AHU-15 AHU15_FIRE_MODE						PHARMA PHARM_AHU14_15,EF3,SB1~2 (700900)		TL122		✓
Pharmacy Rm 6107 FC-601 FCU_601_SCHED						PHARMA PHARM_FCU_601 (700514)		TL9		✓
Pharmacy Unsorted Points CO2_RM_B212						PHARMA PHARM_MISC_BSMNT_AL (700758)		TL3		✓
Pharmacy CRAH System CRAH-2 CRAH2_LO_RT_AL_BV				✓		PHARMA PHARM_HTREJ,HWS,HX3,CRAH (700200)		TL178		✓
Pharmacy Rm 4616 RAD-4-09 RZ4_09_HWRT						PHARMA PHARM_RZ4_04&09 (701340)		TL11		✓
Pharmacy Heating Plant BLR-3 BLR_PB3_S						PHARMA PHARM_BLR1~4,CT,HX1_7,DHW (700800)		TL47		✓
Pharmacy EAV-BE001 BE001_SASH_OPEN_PERCENT						PHARMA PHARM_MACRO_SERVER_FH_TEMP_RM (702900)		TL360		✓
Pharmacy LEF-2 EF-2C EF2_F3_VFD_INST_PWR(kW)						PHARMA PHARM_SB3~6,EF2,SMOKE_EF14~16 (701600)		TL144		✓

elec	elecMeterLoad	enable	energy	entering	enum	equipName	equipRef	esc	exhaust	fan
✓			✓				Pharmacy Heating Plant HX-2 P-HX2A	✓		
✓							Pharmacy Elec Submeters LEED-2N1PC3	✓		
							Pharmacy Rm Corr FC-513	✓		✓
							Pharmacy Rm 1420 FC-111	✓		✓
✓							Pharmacy Elec Submeters LEED-6ETLE1	✓		
							Pharmacy Rm B503 EAV-BS048	✓		
					OFF,ON		Pharmacy AHU-15	✓		
					OFF,ON		Pharmacy Rm 6107 FC-601	✓		
							Pharmacy Unsorted Points	✓		
					OFF,ON		Pharmacy CRAH System CRAH-2	✓		
				✓			Pharmacy Rm 4616 RAD-4-09	✓		
							Pharmacy Heating Plant BLR-3	✓		
							Pharmacy EAV-BE001	✓		
✓							Pharmacy LEF-2 EF-2C	✓		
							Pharmacy LEF-3	✓	✓	
							Pharmacy Unsorted Points	✓	✓	✓



	time	equipRef	groupRef	navName	siteRef	typeRef	unit	value
0	2019-12-13T22:24:37Z	AHU-02	CIRS Air Systems	Discharge Air Temp	CIRS	CIRS_AHU2_SUPPLY_AIR_T	°C	16.707474
1	2019-12-13T22:39:37Z	AHU-02	CIRS Air Systems	Discharge Air Temp	CIRS	CIRS_AHU2_SUPPLY_AIR_T	°C	16.105682
...	...	...	...	...	...	...	...	...
580	2019-12-19T23:39:59Z	AHU-02	CIRS Air Systems	Discharge Air Temp	CIRS	CIRS_AHU2_SUPPLY_AIR_T	°C	23.058758
581	2019-12-19T23:54:59Z	AHU-02	CIRS Air Systems	Discharge Air Temp	CIRS	CIRS_AHU2_SUPPLY_AIR_T	°C	23.049675

Why is the data important?

Knowing end-uses allows UDL to analyze campus wide environmental impact of upgrading specific instruments → reduce energy consumption on campus

Why does UDL need our help?

UDL needs our help because they don't know which instruments are for which end-use. This is because:

1. Data is too granular=hard to pinpoint where the sensor belongs to
2. Data is a bit messy=duplicates & navNames that aren't processed into proper tags

This is where our data cleaning & modelling comes into play...

Energy Consumption (EC) Sensors vs Non-Energy Consumption (NC) Sensors

	equipRef	groupRef	navName	siteRef	typeRef	unit	value
2020-05-31 06:53:17-07:00	Rm 5202 EAV-5E068	Pharmacy Floor 5	Exhaust Air Flow High Lim Sp	Pharmacy	5E068_VLV_FLOW_FDBK_HILIM_SP	L/s	450.000000
2020-05-31 07:08:18-07:00	Rm 6311 EAV-6E049	Pharmacy Floor 6	Exhaust Air Flow High Lim Sp	Pharmacy	6E049_VLV_FLOW_FDBK_HILIM_SP	L/s	250.000000
2020-05-31 09:29:37-07:00	Rm 3335 VAV-3S035	Pharmacy Floor 3	Zone Temp Effective Sp	Pharmacy	VAV_3S035_RT_SP	°C	23.000000
2020-05-31 01:30:00-07:00	Rm 4130 FC-403	Pharmacy Floor 4	Zone Temp Effective Sp	Pharmacy	FCU_403_RT_SP	°C	21.500000
2020-05-31 09:01:23-07:00	Rm 3202 VAV-3S015	Pharmacy Floor 3	Discharge Air Damper Open Cmd	Pharmacy	VAV_3S015_Dmp_Open	omit	True
2020-05-31 10:59:42-07:00	Windows	Pharmacy Floor 5	L5_SE_OAT_CLG_REQUEST	Pharmacy	L5_SE_OAT_CLG_REQUEST	omit	True
2020-05-31 09:38:23-07:00	Elec Submeters LEED-6N4LW1	Pharmacy Utilities	6N4LW1_EnergyPosSum	Pharmacy	6N4LW1_EnergyPosSum	kWh	59165.832031
2020-05-31 09:45:00-07:00	AHU-01 SF	Pharmacy Air Systems	Energy	Pharmacy	AHU1_SF_VFD_PWR(kWh)	kWh	10840.208008
2020-05-31 04:45:00-07:00	Cooling Plant P-9A	Pharmacy Hydronic Systems	Energy	Pharmacy	CHWP_P9A_VFD_PWR(kWh)_TL	kWh	3164.388672
2020-05-31 00:39:54-07:00	Elec Submeters LEED-ATS-S3	Pharmacy Utilities	ATS-S3_EnergyPosSumNR	Pharmacy	ATS-S3_EnergyPosSumNR	kWh	20206020.000000
2020-05-31 04:58:19-07:00	Elec Submeters LEED-ATS-DCB	Pharmacy Utilities	ATS-DCB_EnergyPosSum	Pharmacy	ATS-DCB_EnergyPosSum	kWh	4129881.250000
2020-05-31 04:15:00-07:00	LEF-3 EF-3A	Pharmacy Air Systems	Energy	Pharmacy	EF3_F1_VFD_PWR(kWh)	kWh	43112.464844
2020-05-31 04:28:32-07:00	Elec Submeters LEED-CH-2	Pharmacy Utilities	CH-2_EnergyPosSum	Pharmacy	CH-2_EnergyPosSum	kWh	2470925.500000
2020-05-31 09:30:00-07:00	Cooling Plant P-9B	Pharmacy Hydronic Systems	Energy	Pharmacy	CHWP_P9B_VFD_PWR(kWh)_TL	kWh	3147.771729

NC Data

EC Data

High Level Overview

	equipRef	groupRef	navName	siteRef	typeRef	unit	value
2020-05-31 06:53:17-07:00	Rm 5202 EAV-5E068	Pharmacy Floor 5	Exhaust Air Flow High Lim Sp	Pharmacy	5E068_VLV_FLOW_FDBK_HILIM_SP	L/s	450.000000
2020-05-31 07:08:18-07:00	Rm 6311 EAV-6E049	Pharma			E049_VLV_FLOW_FDBK_HILIM_SP	L/s	250.000000
2020-05-31 09:29:37-07:00	Rm 3335 VAV-3S035	Pharma			VAV_3S035_RT_SP	°C	23.000000
2020-05-31 01:30:00-07:00	Rm 4130 FC-403	Pharma			FCU_403_RT_SP	°C	21.500000
2020-05-31 09:01:23-07:00	Rm 3202 VAV-3S015	Pharma			VAV_3S015_Dmp_Open	omit	True
2020-05-31 10:59:42-07:00	Windows	Pharmacy Floor 5	L5_SE_OAT_CLG_REQUEST	Pharmacy	L5_SE_OAT_CLG_REQUEST	omit	True
2020-05-31 09:38:23-07:00	Elec Submeters LEED-6N4LW1	Pharmacy Utilities	6N4LW1_Energy/PosSum	Pharmacy	6N4LW1_Energy/PosSum	kWh	59165.832031
2020-05-31 09:45:00-07:00	AHU-01 SF	Pharmacy Air Systems	Energy	Pharmacy	AHU1_SF_VFD_PWR(kWh)	kWh	10840.208008
2020-05-31 04:45:00-07:00	Cooling Plant P-9A	Pharmacy Hydronic Systems	Energy	Pharmacy	CHWP_P9A_VFD_PWR(kWh)_TL	kWh	3164.388672
2020-05-31 00:39:54-07:00	Elec Submeters LEED-ATS-S3	Pharmacy Utilities			ATS-S3_Energy/PosSumNR	kWh	20206020.000000
2020-05-31 04:58:19-07:00	Elec Submeters LEED-ATS-DCB	Pharmacy Utilities			ATS-DCB_Energy/PosSum	kWh	4129881.250000
2020-05-31 04:15:00-07:00	LEF-3 EF-3A	Pharmacy Utilities			EF3_F1_VFD_PWR(kWh)	kWh	43112.464844
2020-05-31 04:28:32-07:00	Elec Submeters LEED-CH-2	Pharmacy Utilities	CH-2_Energy/PosSum	Pharmacy	CH-2_Energy/PosSum	kWh	2470925.500000
2020-05-31 09:30:00-07:00	Cooling Plant P-9B	Pharmacy Hydronic Systems	Energy	Pharmacy	CHWP_P9B_VFD_PWR(kWh)_TL	kWh	3147.771729

NC Data

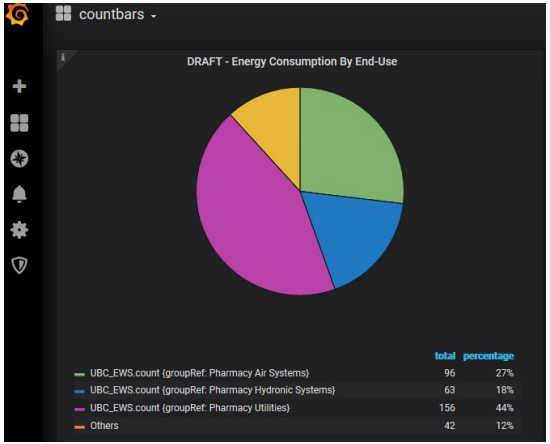
EC Data

Classify EC Data into End-Use Types

(Use NC Data as some of the predictors to help with this)

End-Use
00_HEATING_SPACE_AND_WATER
01_SPACE_COOLING
02_HEATING_COOLING_COMBINED
03_LIGHTING_NORMAL
04_LIGHTING_EMERGENCY
05_OTHER
99_UNKNOWN

Visualize Energy Consumption By End-Use

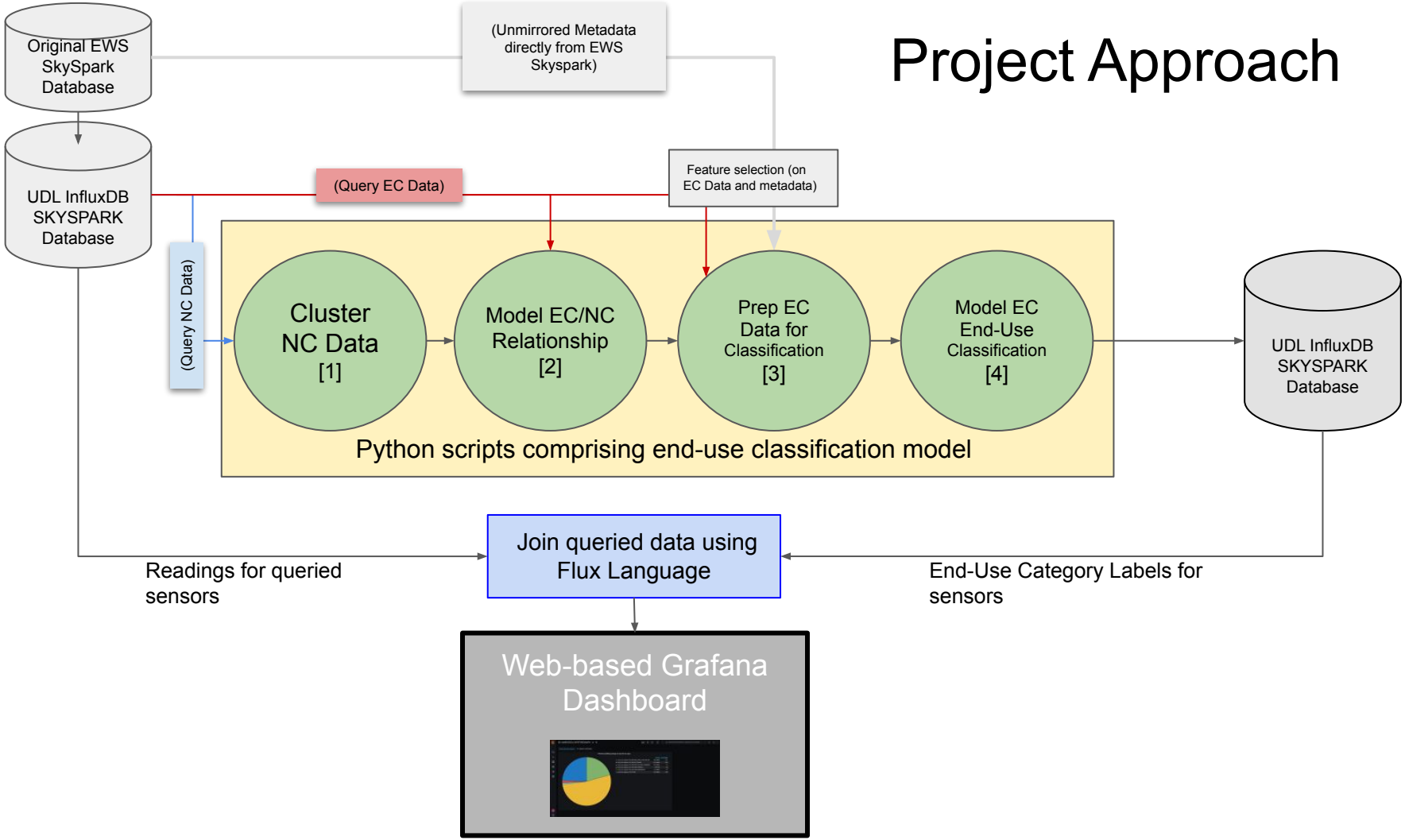


Project Approach

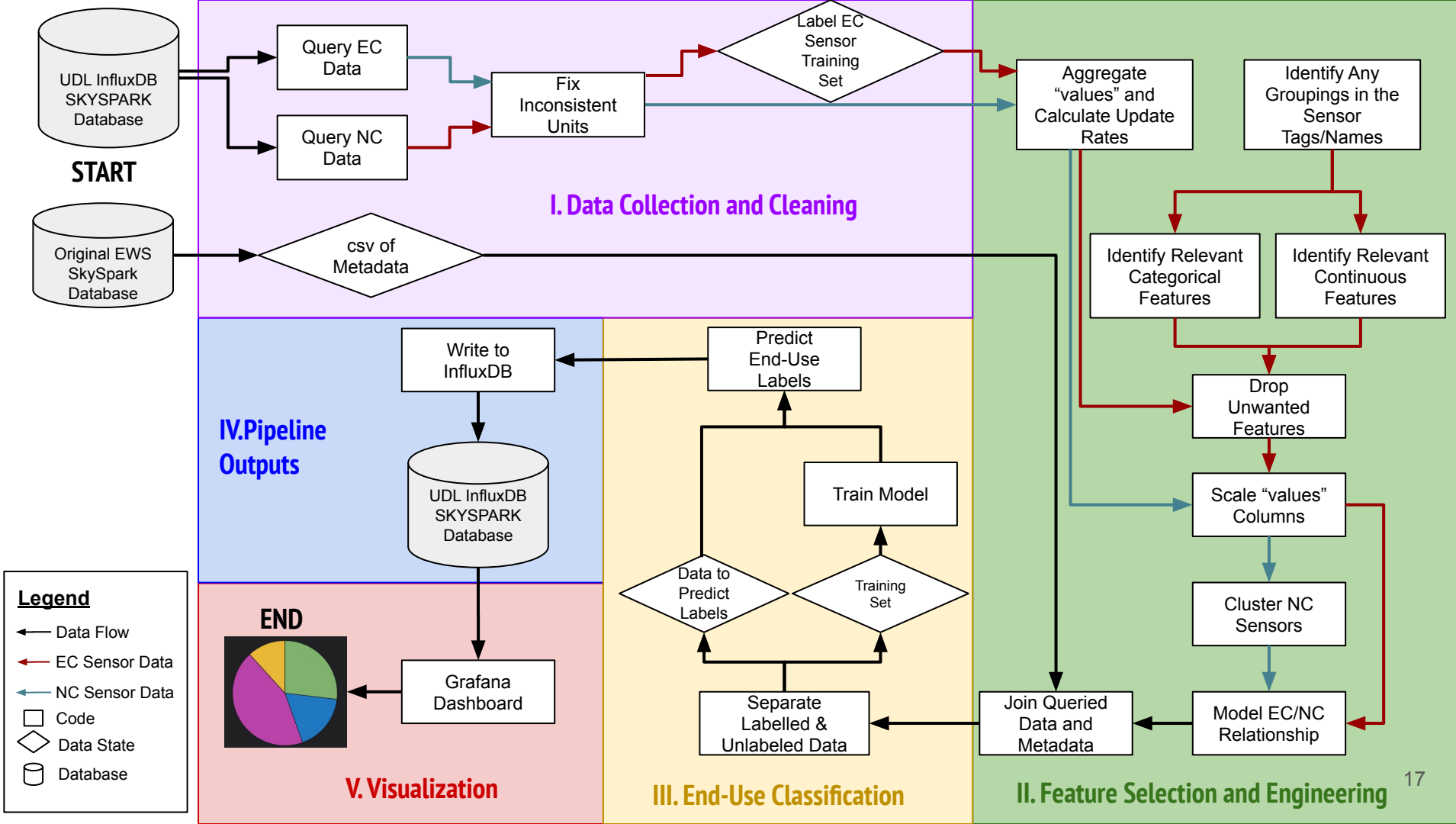
Problem Definition

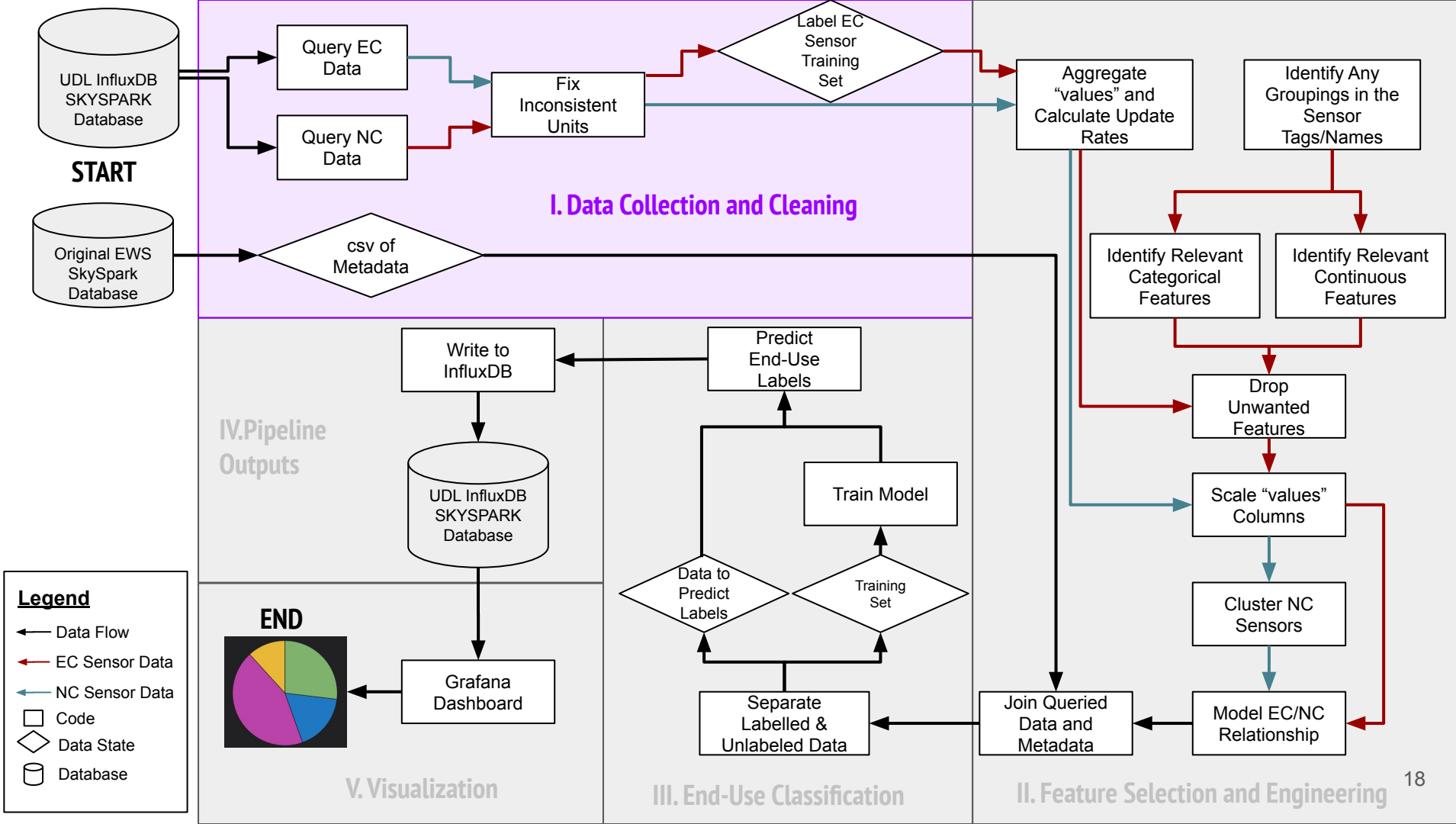
Classify all Pharmacy Building sensors that record energy consumption into appropriate end-uses.

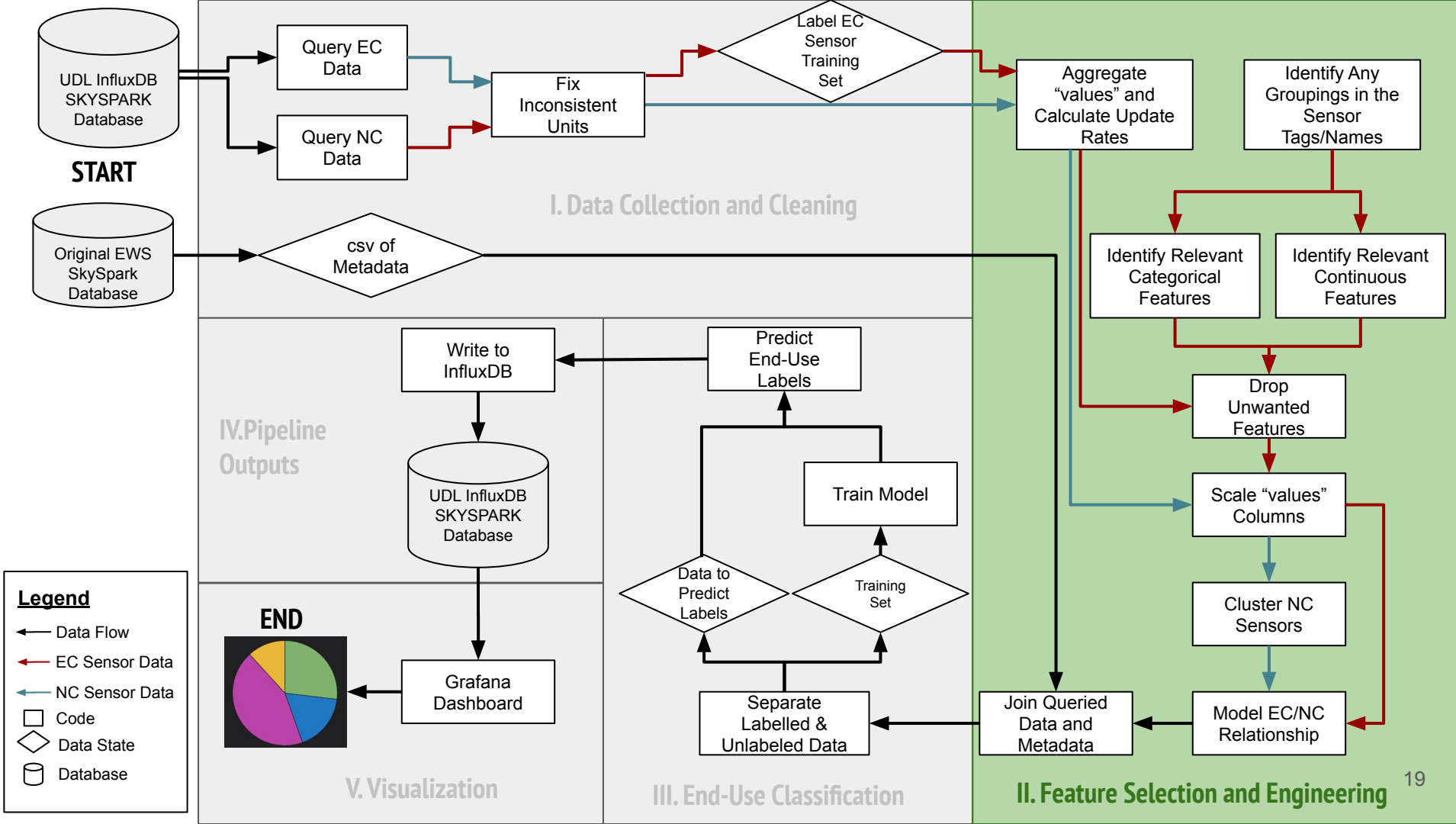
Project Approach

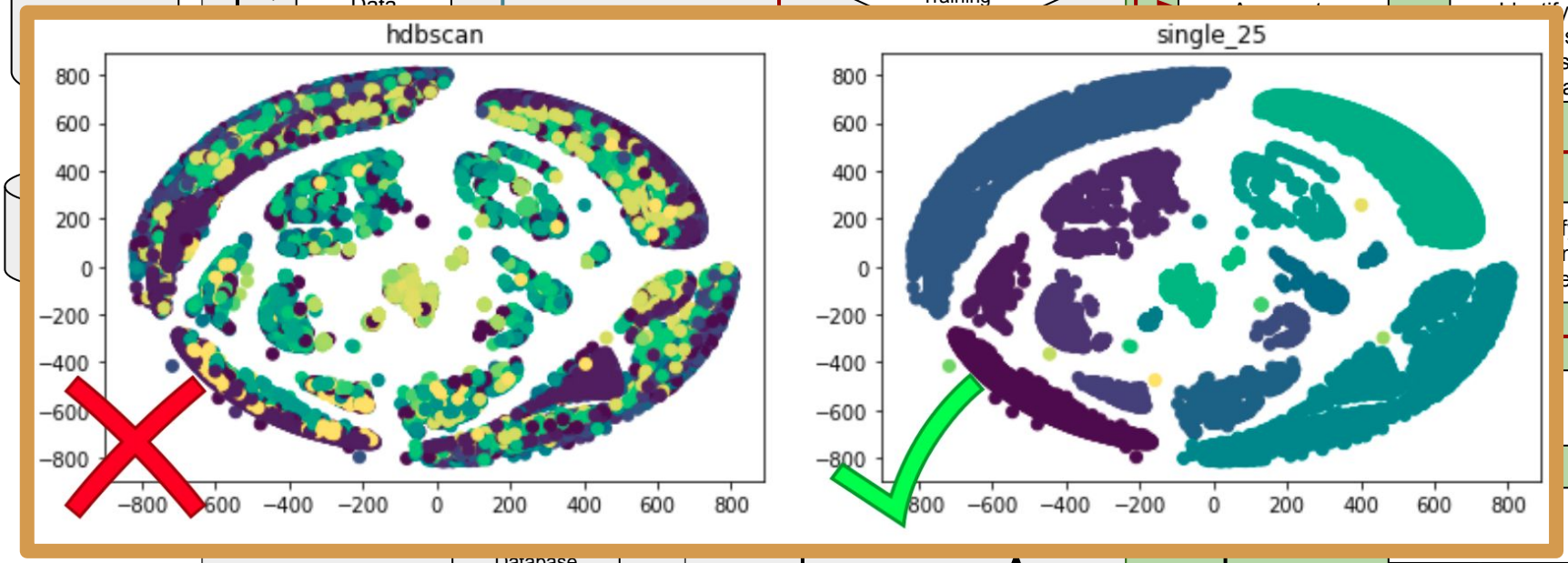


Project Details


UDL InfluxDB SKYSPARK Database
START
Query EC Data
Query NC Data
Fix Inconsistent Units
Label EC Sensor Training Set
Original EWS SkySpark Database
csv of Metadata
I. Data Collection and Cleaning
Aggregate "values" and Calculate Update Rates
Identify Any Groupings in the Sensor Tags/Names
Identify Relevant Categorical Features
Identify Relevant Continuous Features
Drop Unwanted Features
Scale "values" Columns
Cluster NC Sensors
Model EC/NC Relationship
II. Feature Selection and Engineering
Join Queried Data and Metadata
Separate Labelled & Unlabeled Data
Data to Predict Labels
Training Set
Train Model
Predict End-Use Labels
III. End-Use Classification
Write to InfluxDB
UDL InfluxDB SKYSPARK Database
IV.Pipeline Outputs
Grafana Dashboard
END
V. Visualization
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database


UDL InfluxDB SKYSPARK Database
START
Query EC Data
Query NC Data
Fix Inconsistent Units
Label EC Sensor Training Set
I. Data Collection and Cleaning
Original EWS SkySpark Database
csv of Metadata
Aggregate "values" and Calculate Update Rates
Identify Any Groupings in the Sensor Tags/Names
Identify Relevant Categorical Features
Identify Relevant Continuous Features
Drop Unwanted Features
Scale "values" Columns
Cluster NC Sensors
Model EC/NC Relationship
II. Feature Selection and Engineering
Join Queried Data and Metadata
Separate Labelled & Unlabeled Data
Data to Predict Labels
Training Set
Train Model
Predict End-Use Labels
III. End-Use Classification
Write to InfluxDB
UDL InfluxDB SKYSPARK Database
IV.Pipeline Outputs
Grafana Dashboard
END
V. Visualization
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database


UDL InfluxDB SKYSPARK Database
START
Query EC Data
Query NC Data
Fix Inconsistent Units
Label EC Sensor Training Set
I. Data Collection and Cleaning
Original EWS SkySpark Database
csv of Metadata
Aggregate "values" and Calculate Update Rates
Identify Any Groupings in the Sensor Tags/Names
Identify Relevant Categorical Features
Identify Relevant Continuous Features
Drop Unwanted Features
Scale "values" Columns
Cluster NC Sensors
Model EC/NC Relationship
Join Queried Data and Metadata
II. Feature Selection and Engineering
Separate Labelled & Unlabeled Data
Data to Predict Labels
Training Set
Train Model
Predict End-Use Labels
III. End-Use Classification
Write to InfluxDB
UDL InfluxDB SKYSPARK Database
IV.Pipeline Outputs
END
Grafana Dashboard
V. Visualization
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database



Query EC Data

Label EC Sensor Training

Identify Any Gaps in the Sensor Names

Identify Relevant Continuous Features

Database

Cluster NC Sensors

Model EC/NC Relationship

Join Queried Data and Metadata

Separate Labelled & Unlabeled Data

Data to Predict Labels

Training Set

END

Grafana Dashboard

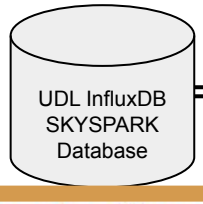
Legend

- Data Flow
- EC Sensor Data
- NC Sensor Data
- Code
- Data State
- Database

V. Visualization

III. End-Use Classification

II. Feature Selection and Engineering



Query EC Data

Query NC

Fix Inconsistent Units

Label EC Sensor Training Set

Aggregate "values" and Calculate Update Rates

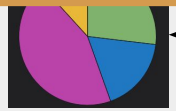
Identify Any Groupings in the Sensor Tags/Names

Equation Generated from Ridge Regression for each sensor

$$ec_sensor_response = \beta_0 + \beta_1 mean_{c1} + \beta_2 max_{c1} + \beta_3 min_{c1} + \beta_4 std_{c1} + \beta_5 u_rate_{c1} + \ldots + \beta_{5*n-4} mean_{cn} + \beta_{5*n-3} max_{cn} + \beta_{5*n-2} min_{cn} + \beta_{5*n-1} std_{cn} + \beta_{5*n} u_rate_{cn}$$

0	1	2	3	...	5n-2	5n-1	5n	uniqueID
0.000037	-0.004377	0.0	-0.000041	...	5.876493	8.502804	20.087383	AHU-01 SF Air Systems Energy AHU1_SF_VFD_PWR(kWh)
0.000039	-0.004622	0.0	-0.000044	...	6.537176	8.851925	20.473544	AHU-02 SF Air Systems Energy AHU2_SF_VFD_PWR(kWh)

NC Sensor Data
Code
Data State
Database



Grafana Dashboard

Separate Labelled & Unlabeled Data

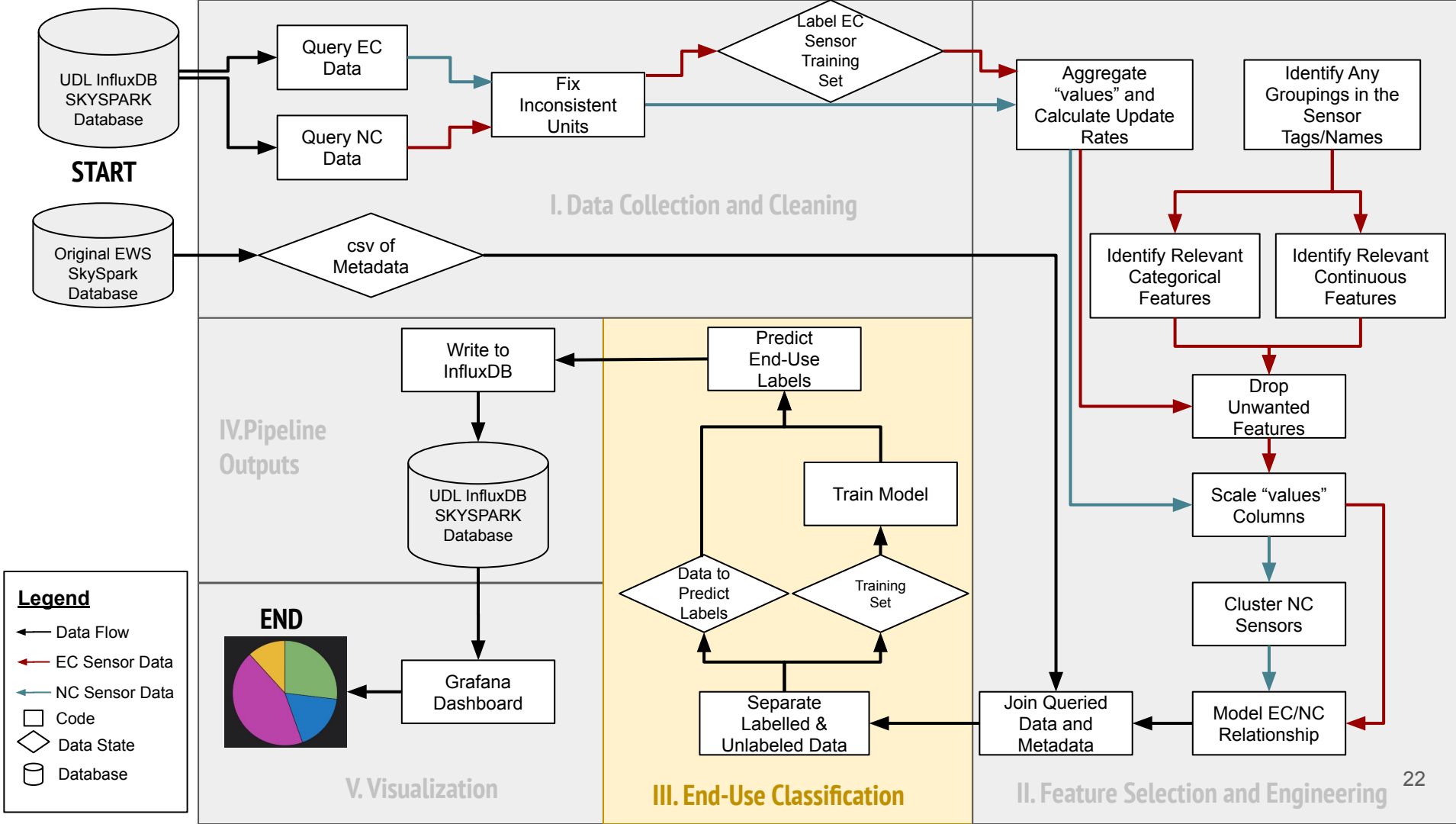
Join Queried Data and Metadata

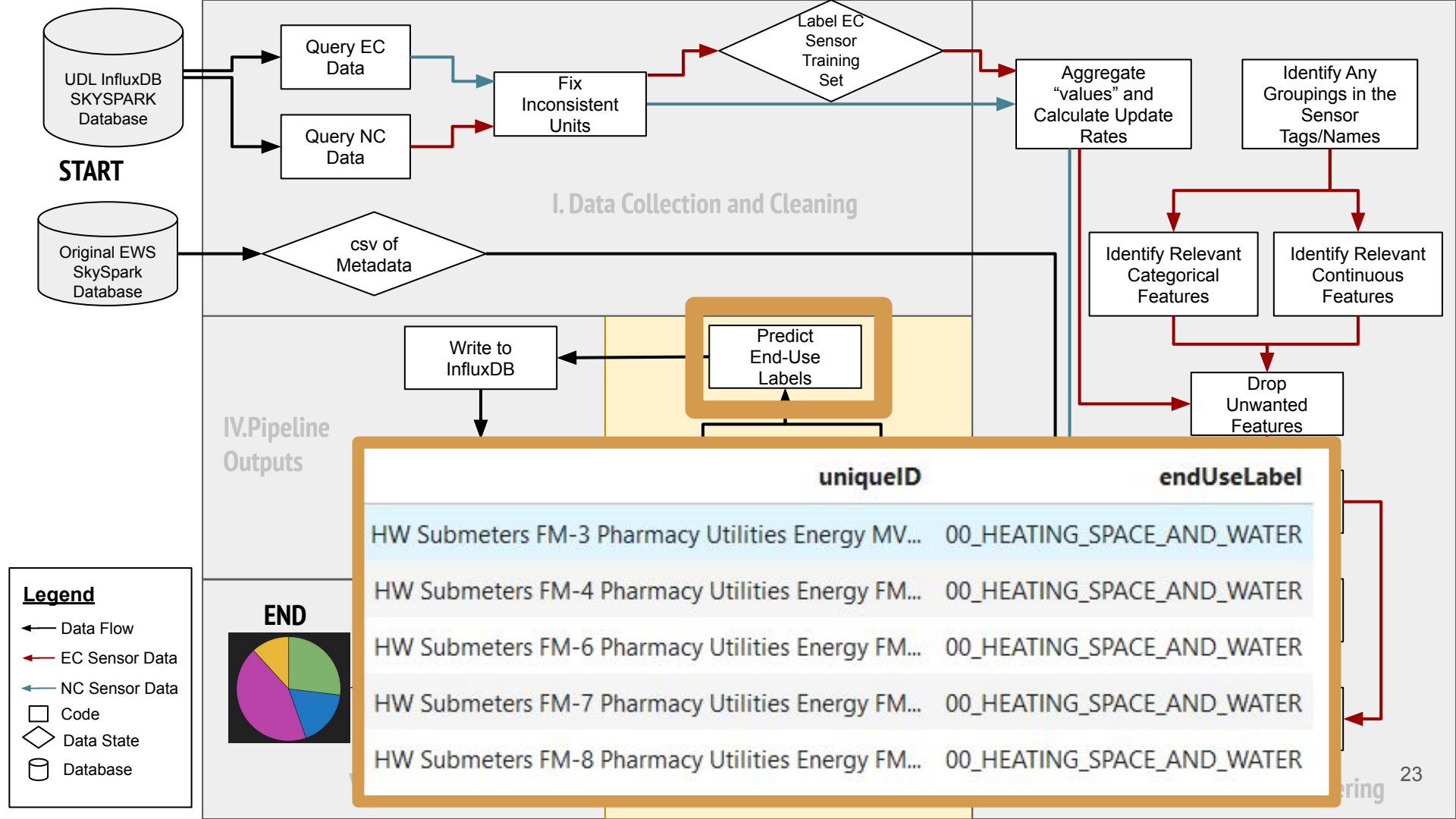
Model EC/NC Relationship

V. Visualization

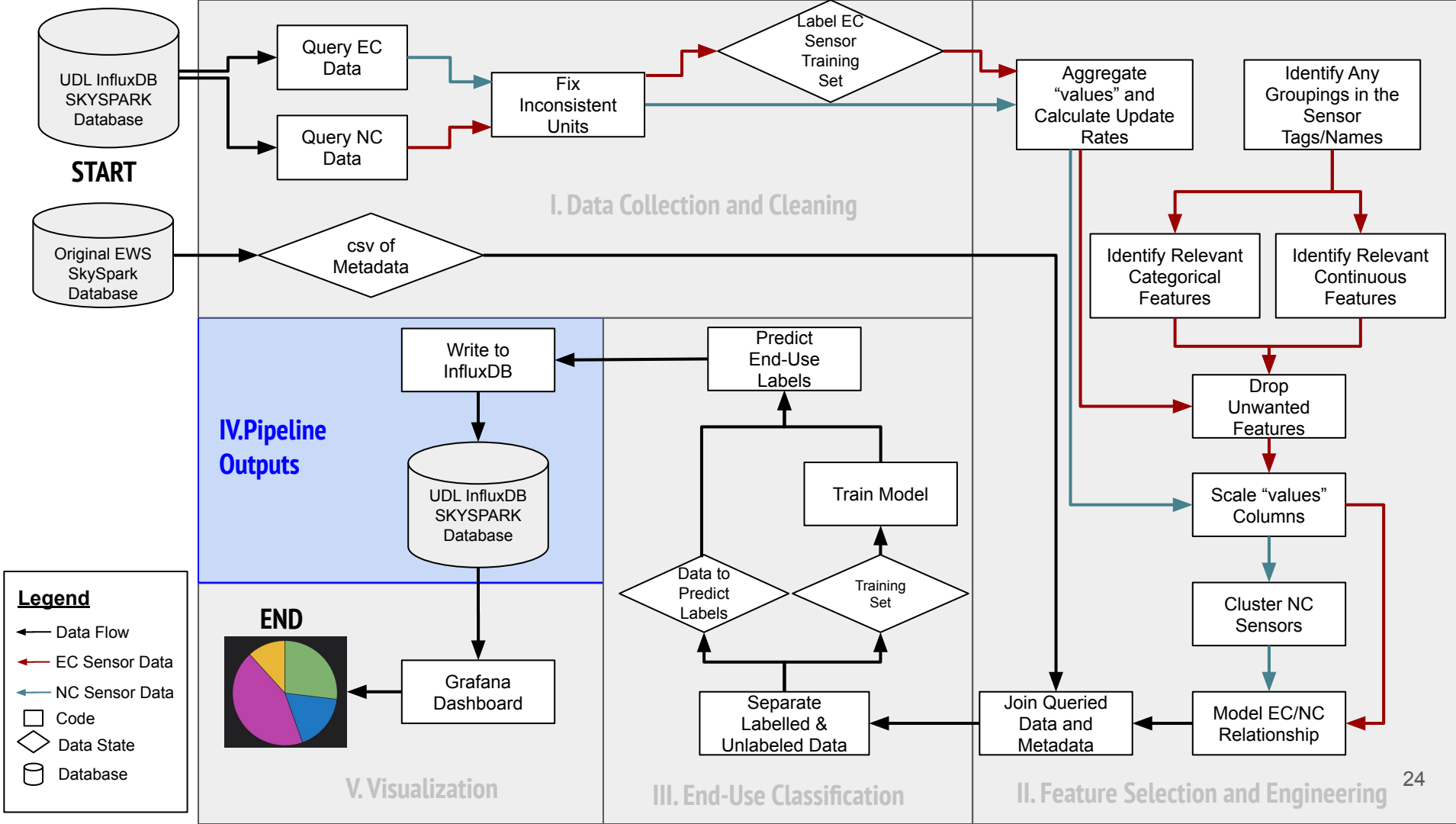
III. End-Use Classification

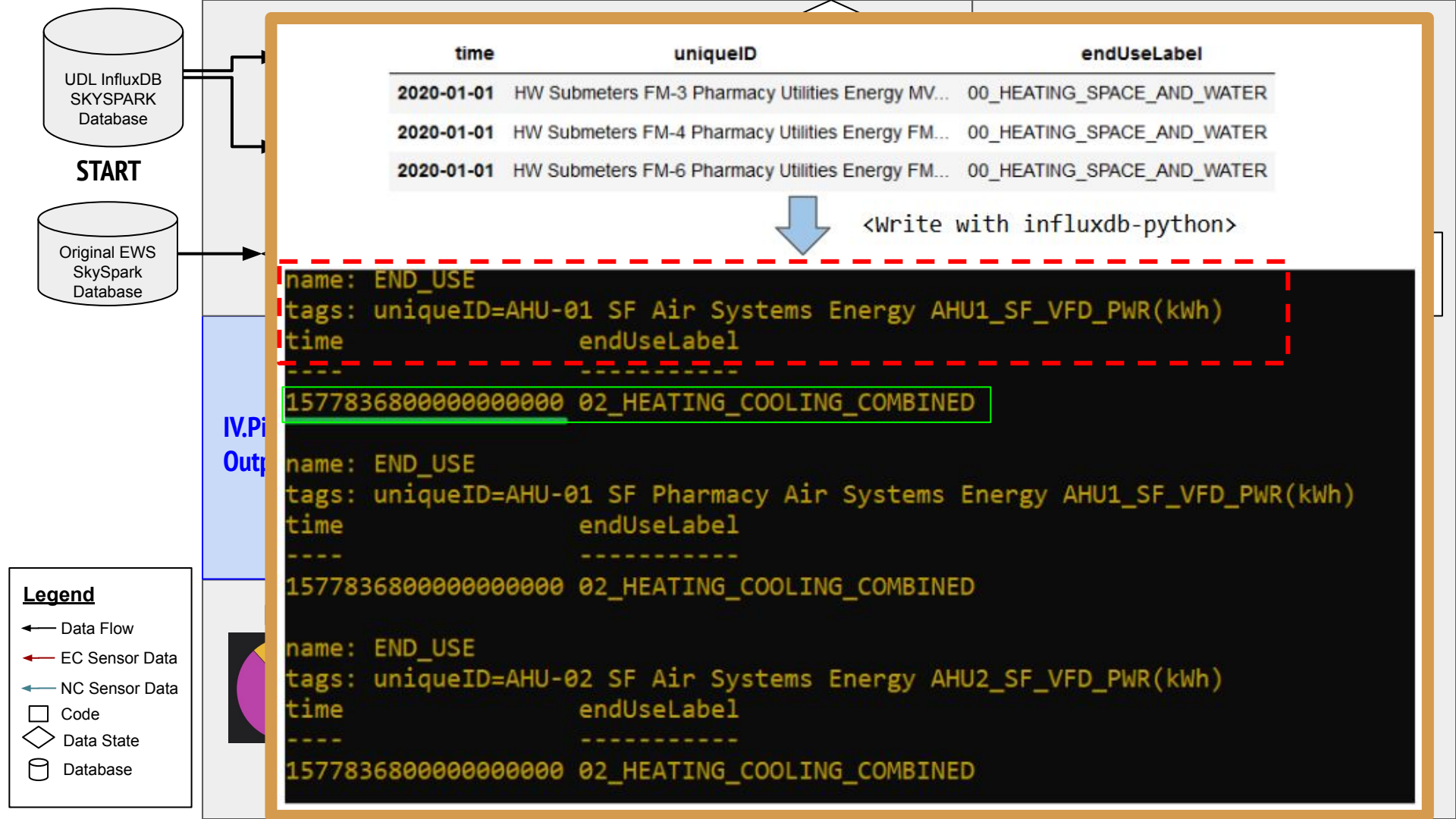
II. Feature Selection and Engineering

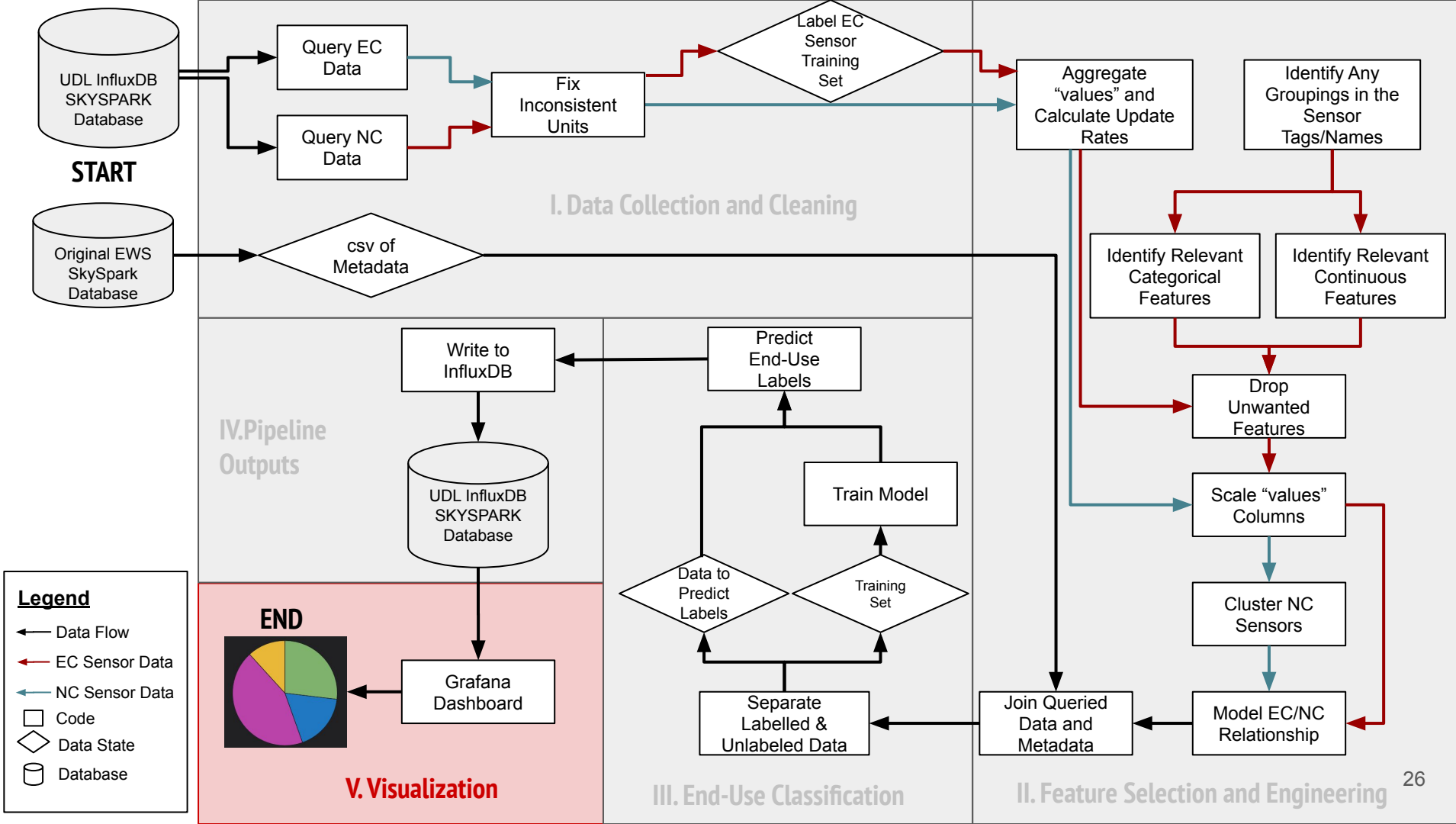

UDL InfluxDB SKYSPARK Database
START
Query EC Data
Query NC Data
Fix Inconsistent Units
Label EC Sensor Training Set
I. Data Collection and Cleaning
Original EWS SkySpark Database
csv of Metadata
Aggregate "values" and Calculate Update Rates
Identify Any Groupings in the Sensor Tags/Names
Identify Relevant Categorical Features
Identify Relevant Continuous Features
Drop Unwanted Features
Scale "values" Columns
Cluster NC Sensors
Model EC/NC Relationship
Join Queried Data and Metadata
II. Feature Selection and Engineering
Separate Labelled & Unlabeled Data
Data to Predict Labels
Training Set
Train Model
Predict End-Use Labels
III. End-Use Classification
Write to InfluxDB
UDL InfluxDB SKYSPARK Database
IV.Pipeline Outputs
Grafana Dashboard
END
V. Visualization
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database

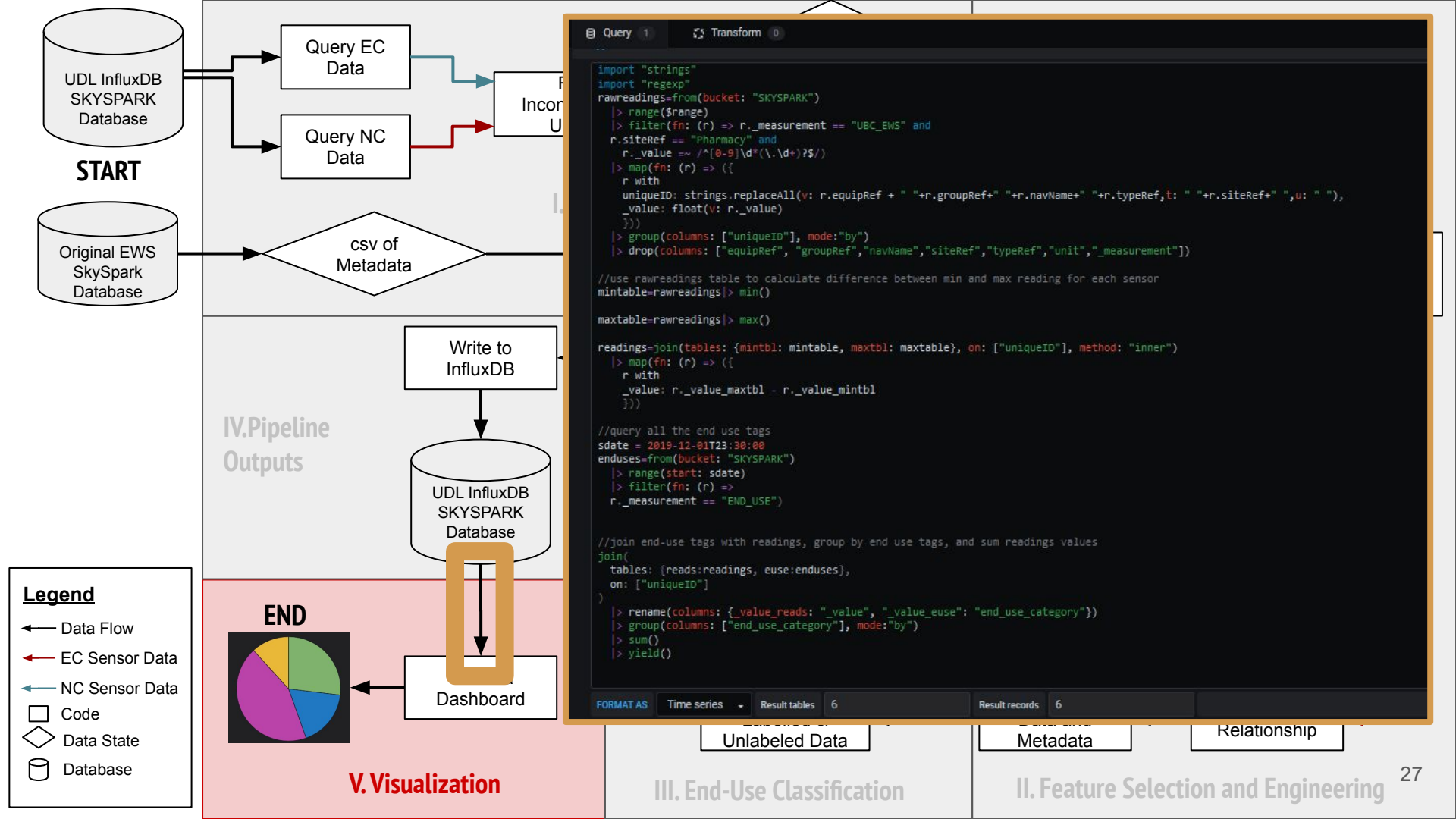


uniqueID	endUseLabel
HW Submeters FM-3 Pharmacy Utilities Energy MV...	00_HEATING_SPACE_AND_WATER
HW Submeters FM-4 Pharmacy Utilities Energy FM...	00_HEATING_SPACE_AND_WATER
HW Submeters FM-6 Pharmacy Utilities Energy FM...	00_HEATING_SPACE_AND_WATER
HW Submeters FM-7 Pharmacy Utilities Energy FM...	00_HEATING_SPACE_AND_WATER
HW Submeters FM-8 Pharmacy Utilities Energy FM...	00_HEATING_SPACE_AND_WATER


UDL InfluxDB SKYSPARK Database
START
Query EC Data
Query NC Data
Fix Inconsistent Units
Label EC Sensor Training Set
Aggregate "values" and Calculate Update Rates
Identify Any Groupings in the Sensor Tags/Names
I. Data Collection and Cleaning
Original EWS SkySpark Database
csv of Metadata
Identify Relevant Categorical Features
Identify Relevant Continuous Features
Drop Unwanted Features
Scale "values" Columns
Cluster NC Sensors
Model EC/NC Relationship
Join Queried Data and Metadata
II. Feature Selection and Engineering
Separate Labelled & Unlabeled Data
Data to Predict Labels
Training Set
Train Model
Predict End-Use Labels
III. End-Use Classification
IV.Pipeline Outputs
Write to InfluxDB
UDL InfluxDB SKYSPARK Database
END
Grafana Dashboard
V. Visualization
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database


UDL InfluxDB SKYSPARK Database
START
Original EWS SkySpark Database
IV.Pi
Outp
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database
time
uniqueID
endUseLabel
2020-01-01 HW Submeters FM-3 Pharmacy Utilities Energy MV... 00_HEATING_SPACE_AND_WATER
2020-01-01 HW Submeters FM-4 Pharmacy Utilities Energy FM... 00_HEATING_SPACE_AND_WATER
2020-01-01 HW Submeters FM-6 Pharmacy Utilities Energy FM... 00_HEATING_SPACE_AND_WATER
<Write with influxdb-python>
name: END_USE
tags: uniqueID=AHU-01 SF Air Systems Energy AHU1_SF_VFD_PWR(kWh)
time endUseLabel
---- -----------
1577836800000000000 02_HEATING_COOLING_COMBINED
name: END_USE
tags: uniqueID=AHU-01 SF Pharmacy Air Systems Energy AHU1_SF_VFD_PWR(kWh)
time endUseLabel
---- -----------
1577836800000000000 02_HEATING_COOLING_COMBINED
name: END_USE
tags: uniqueID=AHU-02 SF Air Systems Energy AHU2_SF_VFD_PWR(kWh)
time endUseLabel
---- -----------
1577836800000000000 02_HEATING_COOLING_COMBINED


UDL InfluxDB SKYSPARK Database
START
Query EC Data
Query NC Data
Fix Inconsistent Units
Label EC Sensor Training Set
Original EWS SkySpark Database
csv of Metadata
I. Data Collection and Cleaning
Aggregate "values" and Calculate Update Rates
Identify Any Groupings in the Sensor Tags/Names
Identify Relevant Categorical Features
Identify Relevant Continuous Features
Drop Unwanted Features
Scale "values" Columns
Cluster NC Sensors
Model EC/NC Relationship
II. Feature Selection and Engineering
Join Queried Data and Metadata
Separate Labelled & Unlabeled Data
Data to Predict Labels
Training Set
Train Model
Predict End-Use Labels
III. End-Use Classification
Write to InfluxDB
UDL InfluxDB SKYSPARK Database
IV.Pipeline Outputs
Grafana Dashboard
END
V. Visualization
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database


UDL InfluxDB SKYSPARK Database
Query EC Data
Query NC Data
START
Original EWS SkySpark Database
csv of Metadata
Write to InfluxDB
IV.Pipeline Outputs
UDL InfluxDB SKYSPARK Database
END
Dashboard
Legend
Data Flow
EC Sensor Data
NC Sensor Data
Code
Data State
Database
V. Visualization
Unlabeled Data
III. End-Use Classification
Metadata
Relationship
II. Feature Selection and Engineering
Query
Transform
import "strings"
import "regexp"
rawreadings=from(bucket: "SKYSPARK")
|> range($range)
|> filter(fn: (r) => r._measurement == "UBC_EWS" and
r.siteRef == "Pharmacy" and
r._value =~ /^[0-9]\d*(\.\d+)?$/)
|> map(fn: (r) => ({
r with
uniqueID: strings.replaceAll(v: r.equipRef + " "+r.groupRef+" "+r.navName+" "+r.typeRef,t: " "+r.siteRef+" ",u: " "),
_value: float(v: r._value)
}))
|> group(columns: ["uniqueID"], mode:"by")
|> drop(columns: ["equipRef", "groupRef","navName","siteRef","typeRef","unit","_measurement"])
//use rawreadings table to calculate difference between min and max reading for each sensor
mintable=rawreadings|> min()
maxtable=rawreadings|> max()
readings=join(tables: {mintbl: mintable, maxtbl: maxtable}, on: ["uniqueID"], method: "inner")
|> map(fn: (r) => ({
r with
_value: r._value_maxtbl - r._value_mintbl
}))
//query all the end use tags
sdate = 2019-12-01T23:30:00
enduses=from(bucket: "SKYSPARK")
|> range(start: sdate)
|> filter(fn: (r) =>
r._measurement == "END_USE")
//join end-use tags with readings, group by end use tags, and sum readings values
join(
tables: {reads:readings, euse:enduses},
on: ["uniqueID"]
)
|> rename(columns: {_value_reads: "_value", "_value_euse": "end_use_category"})
|> group(columns: ["end_use_category"], mode:"by")
|> sum()
|> yield()
FORMAT AS
Time series
Result tables
6
Result records
6

Results & Analysis

Model Comparison

→ Predicting End-Use Labels

Clustering Model	Linkage Method	Number of Clusters	Supervised Model	accuracy	precision	recall	f1_score	log loss
Agglomerative	Average	15	Bagging	0.9143	0.9310	0.9143	0.9144	0.2683
Agglomerative	Complete	16	Random Forest	0.9429	0.9524	0.9429	0.9449	0.3786
K-Means	-	15	Extremely Random Trees	0.9143	0.9310	0.9143	0.9144	1.2282
Agglomerative	Single	20	Bagging	0.9429	0.9524	0.9429	0.9449	0.1961
Agglomerative	Single	25	Gradient Boost	0.9429	0.9524	0.9429	0.9449	0.1979
VBGM	-	19	Gradient Boost	0.9429	0.9490	0.9429	0.9389	0.4589
Agglomerative	Ward	15	Bagging	0.9143	0.9310	0.9143	0.9144	0.3518

Confusion Matrix for Single Linkage with 20 Clusters

	00	01	02	03	04	05
00_HEATING_SPACE_AND_WATER	[10	0	0	0	0	0]
01_SPACE_COOLING	[0	5	0	0	0	0]
02_HEATING_COOLING_COMBINED	[0	0	8	0	0	1]
03_LIGHTING_NORMAL	[0	0	0	2	0	0]
04_LIGHTING_EMERGENCY	[0	0	0	0	2	0]
05_OTHER	[0	0	0	1	0	6]

Final Results

End Use Category	Sensor Count	% of Sensors
00_HEATING_SPACE_AND_WATER	54	26%
01_SPACE_COOLING	35	17%
02_HEATING_COOLING_COMBINED	39	19%
03_LIGHTING_NORMAL	26	13%
04_LIGHTING_EMERGENCY	10	5%
05_OTHER	44	21%
Total	**208**	**100%**

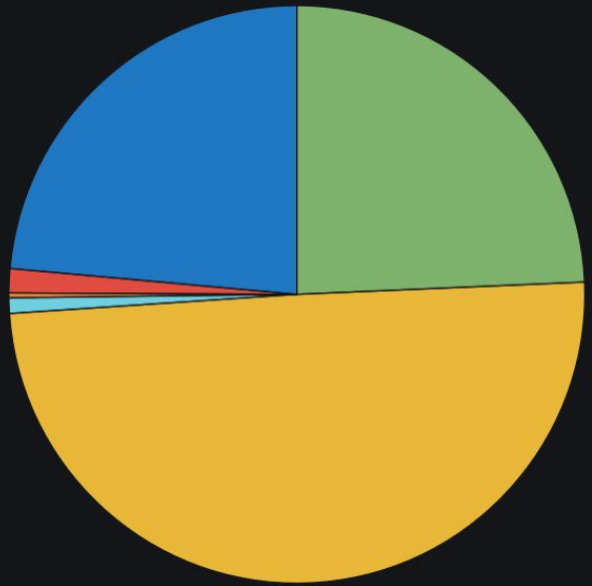
Dashboard Demo

Pie_For_Presentation

2020-05-27 00:00:00 to 2020-05-28 23:59:59

Choose End use Category: Pharmacy

Pharmacy Building Energy use by end-use type



	current	percentage
end_use_category=00_HEATING_SPACE_AND_WATER	116.7 MWh	24%
end_use_category=01_SPACE_COOLING	238 MWh	50%
end_use_category=02_HEATING_COOLING_COMBINED	4.16 MWh	1%
end_use_category=03_LIGHTING_NORMAL	1.294 MWh	0%
end_use_category=04_LIGHTING_EMERGENCY	6.48 MWh	1%
end_use_category=05_OTHER	113.1 MWh	24%

Conclusion

Conclusion

Final Model:

Clustering Model	Linkage Method	Number of Clusters	Supervised Model
Agglomerative	Single	20	Bagging

Model Accuracy: 94.29%

Model Precision: 95.24%

Difficulties & Lessons Learned

- Initially thought all sensors needed to be assigned an end-use label → redesign of the planned model

- Took lots of time to understand the data well enough to label it for a training & test set

- Assorted InfluxDB-related issues (i.e. all readings stored as strings instead of float data-type, unique identifier changes over time for sensor, etc)-> time spent finding fixes/work-arounds.

Future Direction of Work

01	**Accuracy Updates**	• Include more UBC buildings -> bigger training set • Modify code to work with updated database structure -> uniqueID+more days of data will improve accuracy • Additional feature engineering like power consumption curve shapes • Investigate+label hierarchy of meters in Pharmacy building → fix issue with multi-counting same energy use in pie-chart
02	**Performance Updates**	• Take advantage of new database structure to aggregate EC/NC data in queries and simplify flux query for visualization. • Misc. optimization of code - > be able to handle more data
03	**Scalability Updates**	• Build model for meter hierarchy classification (mentioned above) • Make Feature Selection code dynamic • Update code to work with multiple buildings

Acknowledgements

- Jiachen Wei & Mike Kennedy from UDL
- Scott Fazackerley, Jeff Bulmer, Debangsha Sarkar from UBCO
- Jeff Andrews from UBCO

Questions