Accident-Weather Analysis Tool

DATA 515 Final Project

Spring 2017

Michael Grant, Dane Jordan, Samir Patel, Rex Thompson

WEATHER UNDERGROUND

WASHINGTON STATE PATROL

Problem Statement

- ~42,000 collisions reported in the state in the last year
- Weather is a contributing factor even if not the primary cause
- Lack of objective quantitative collision-specific weather data
- Washington State Patrol (WSP) maintains detailed collision database; however, weather observation quality is low:
  - Observations made by untrained weather professionals
  - Observations are qualitative, not quantitative
- High-quality weather data available, but sparse
  - Road Weather Information Service (RWIS)
  - Official observation stations (e.g. airports)

Wea_Typ_Cd	Descr
0	Unknown
1	Clear or Partly Cloudy
2	Overcast
3	Raining
4	Snowing
5	Fog or Smog or Smoke
6	Sleet or Hail or Freezing Rain
7	Severe Crosswind
8	Blowing Sand or Dirt or Snow
9	Other

Proposed Solution

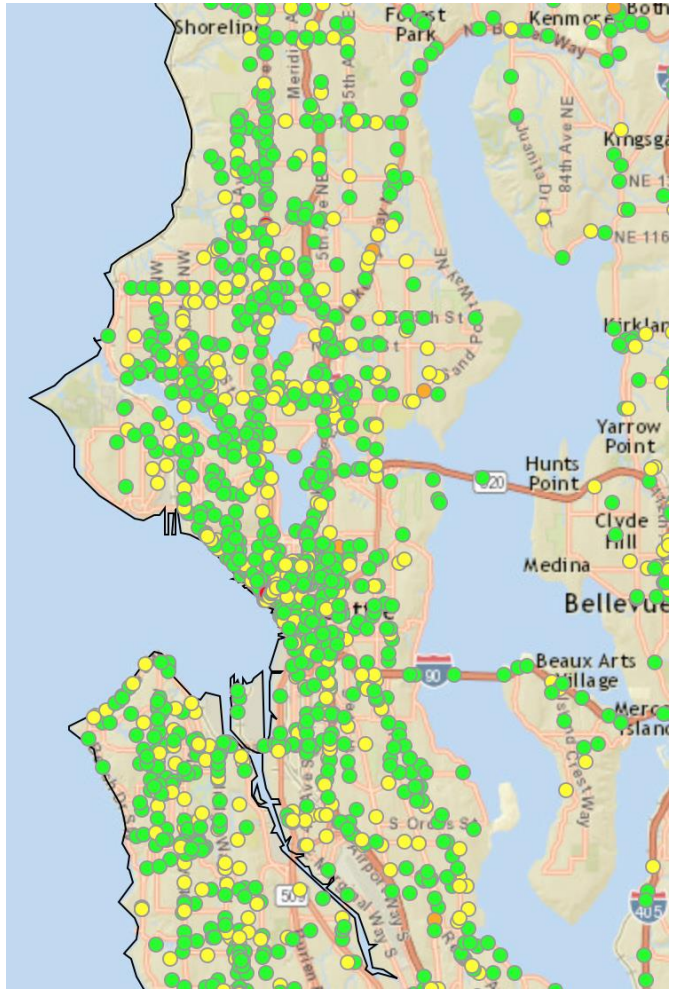
- Weather Underground Personal Weather Station (PWS) network
  - Good geographic coverage, especially in populated well-to-do areas
  - Objective time- and location-specific quantitative weather data
  - Data quality not guaranteed, but probably “close enough” to paint a picture of conditions leading up to and during a collision
- **GOAL**: Create a tool to enable enhancement of the WSP collision database with objective observations from nearby personal weather stations

Target Users

- Traffic Engineers
  - Weather-dependent traffic control measures
  - Long-term strategic planning
- Law Enforcement Officers
  - Proactive safety measures
  - Patrol location adjustment
- Insurance Companies
  - Premium tuning
  - Claims adjustment
- Common Driver
  - Map integration

Washington State Patrol (WSP) Collision Analysis Database

- 14 tables with 217 fields
- 42,000 rows from May 2016 – April 2017
- Updated to-the-minute by law enforcement
- Key features:
  - Date and time
  - Location
  - Collision type
  - Basic weather information
  - Much, much more…

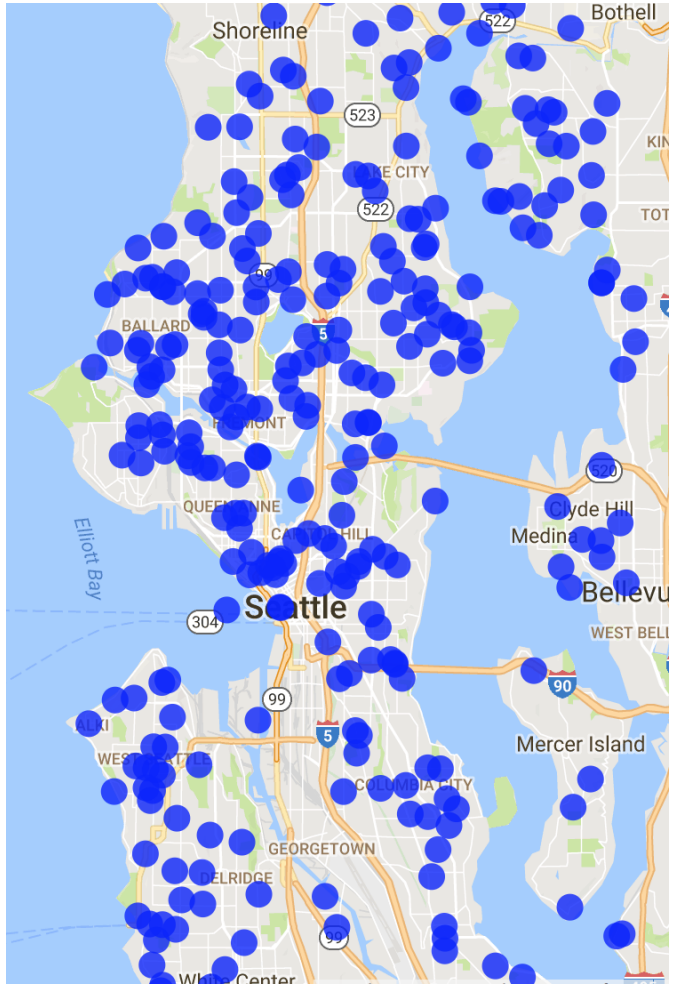


https://fortress.wa.gov/wsp/collisionanalysistool/

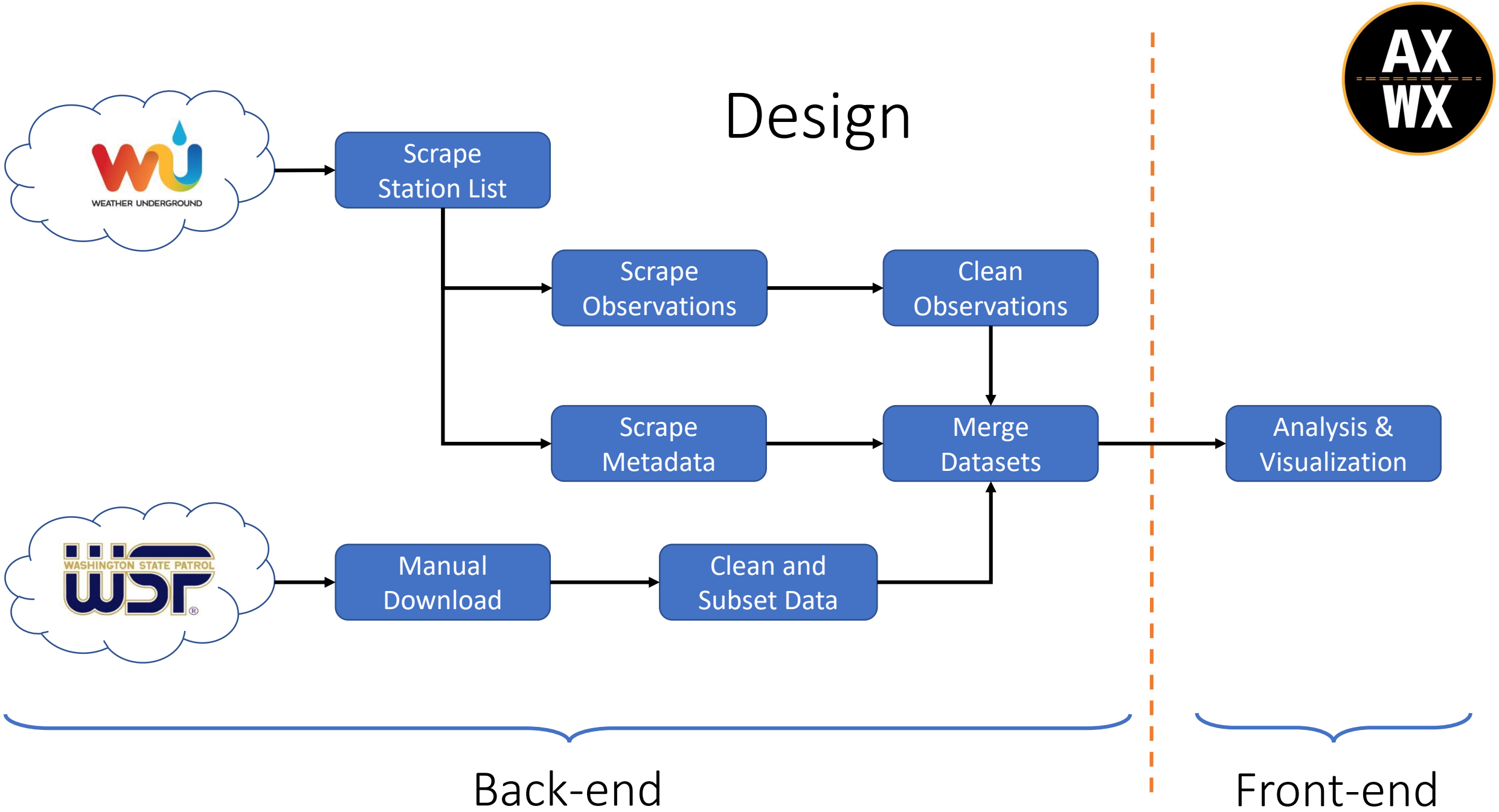
Weather Underground Personal Weather Station (PWS) Network

- 365+ stations in greater Seattle area
- Most stations report in 5-minute intervals
- Key features:
  - Temperature
  - Relative Humidity
  - Wind Speed/Direction/Gusts
  - Precipitation
  - Pressure

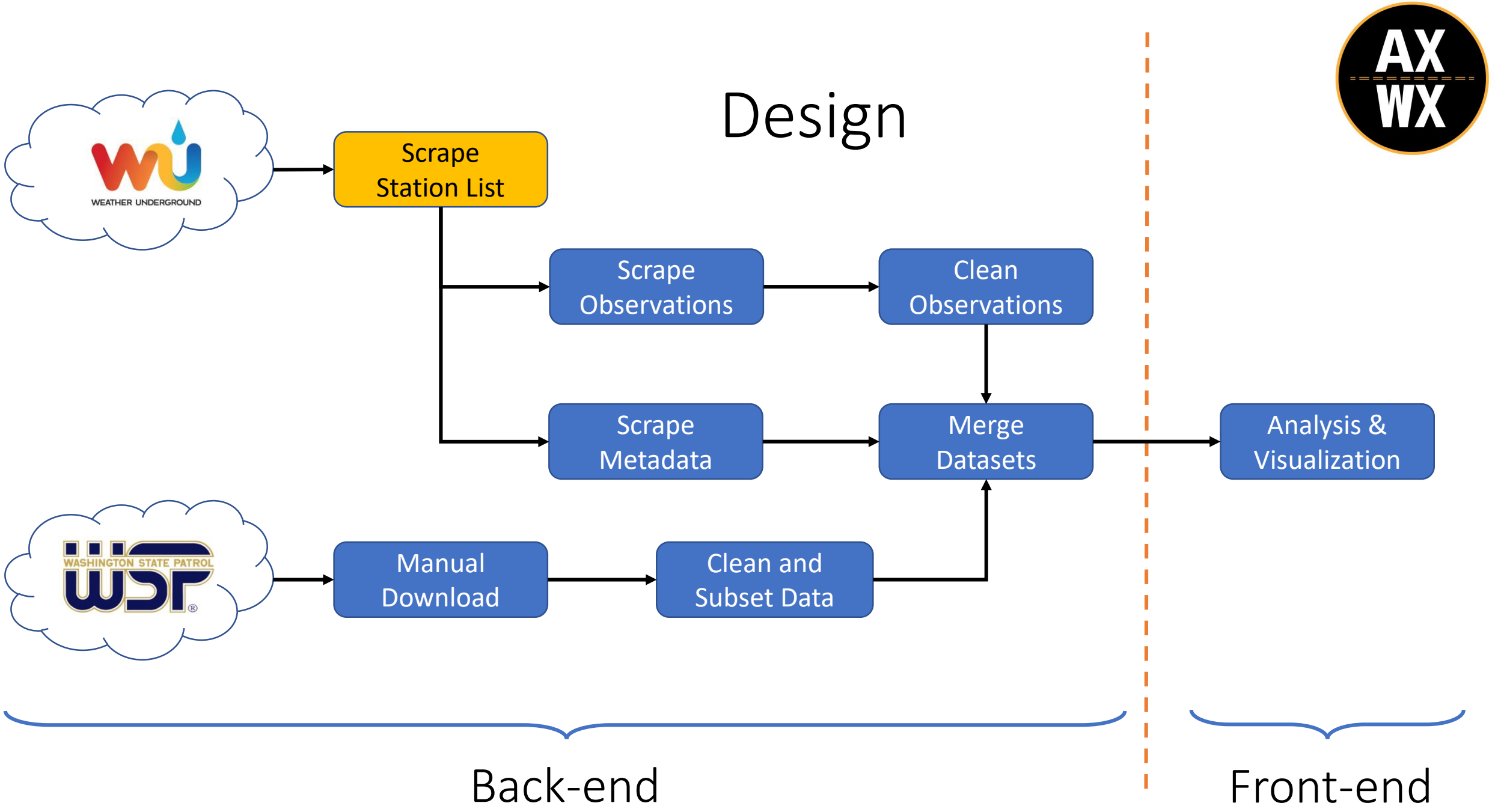
https://www.wunderground.com/weatherstation/ListStations.asp?selectedState=WA&selectedCountry=United+States



Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

Weather Stations in WA, United States

Station ID	Neighborhood	City	Station Type	Site
KWAVENER2	Alpine, WA	0.8Mi NNE of Venersborg Store, Venersborg	Davis Vantage Vue	
KWAMARYS7	Kruse Junction	North Marysville	Davis Vantage Pro 2	
KWAWALLA2	Foothills of the Blue Mountains	Walla Walla	Davis Vantage Pro2 (Wireless)	
KWAABERD14	Think of Me Hill	Aberdeen	Davis Vantage Pro2 Plus (Wireless)	
KWAABERD15	Central Park	Aberdeen	AcuRite Pro Weather Center	
KWAABERD10	Aberdeen below Hospital	Aberdeen	AcuRite Pro Weather Center	
KWAABERD13	Sumner & Myrtle, Aberdeen	Aberdeen	RainWise MK-III-LR	
KWAABERD16	Pacific Ave	Aberdeen	AcuRite Pro Weather Center	
KWAABERD11	Central Park - Crash's Yard	Aberdeen	Davis Vantage Pro2 (Wireless)	
KWAADAMS2	Seattle	Adams	Netatmo	
KWAALLYN3	Twanoh	Allyn	Ambient Weather WS-1001-WiFi (Wireless)	
KWAALLYN4	Allyn View Estates	Allyn	AcuRite Pro Weather Center	
KWAALLYN2	Grapeview	Allyn-Grapeview	Ambient Weather WS-1001-WiFi (Wireless)	
KWAAMBOY5	.3 miles NW of Saddle Dam	Amboy	Davis Vantage Pro2 (Wireless)	
KWAAMBOY10	Grantham Amboy	Amboy	AcuRite Pro Weather Center	

Browse Our PWS Network

250,000+ personal weather stations across the globe already send data to Weather Underground! Browse the PWS contributing to our network:

Select a state View

Select a country View

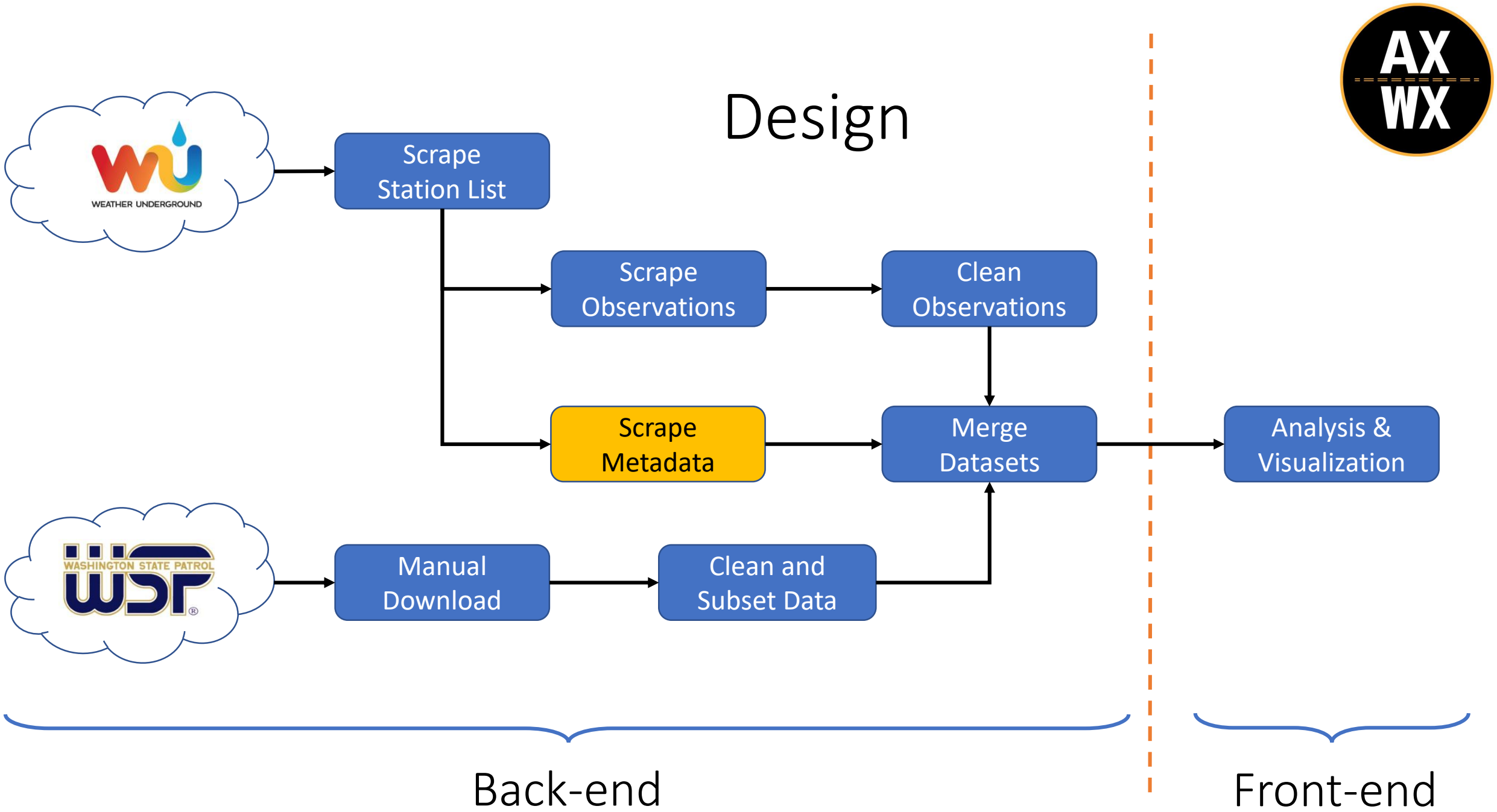
WunderWiki Help Topics

If you require assistance, or would like to learn more about Personal Weather Stations, please select a related topic below:

- About Personal Weather Stations
- Software Configuration
- Frequently Asked Questions
- Rapid Fire Updates
- Other Website Resources
- Upload Protocol (for programmers)

Desktop Applications

Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

https://api.wunderground.com

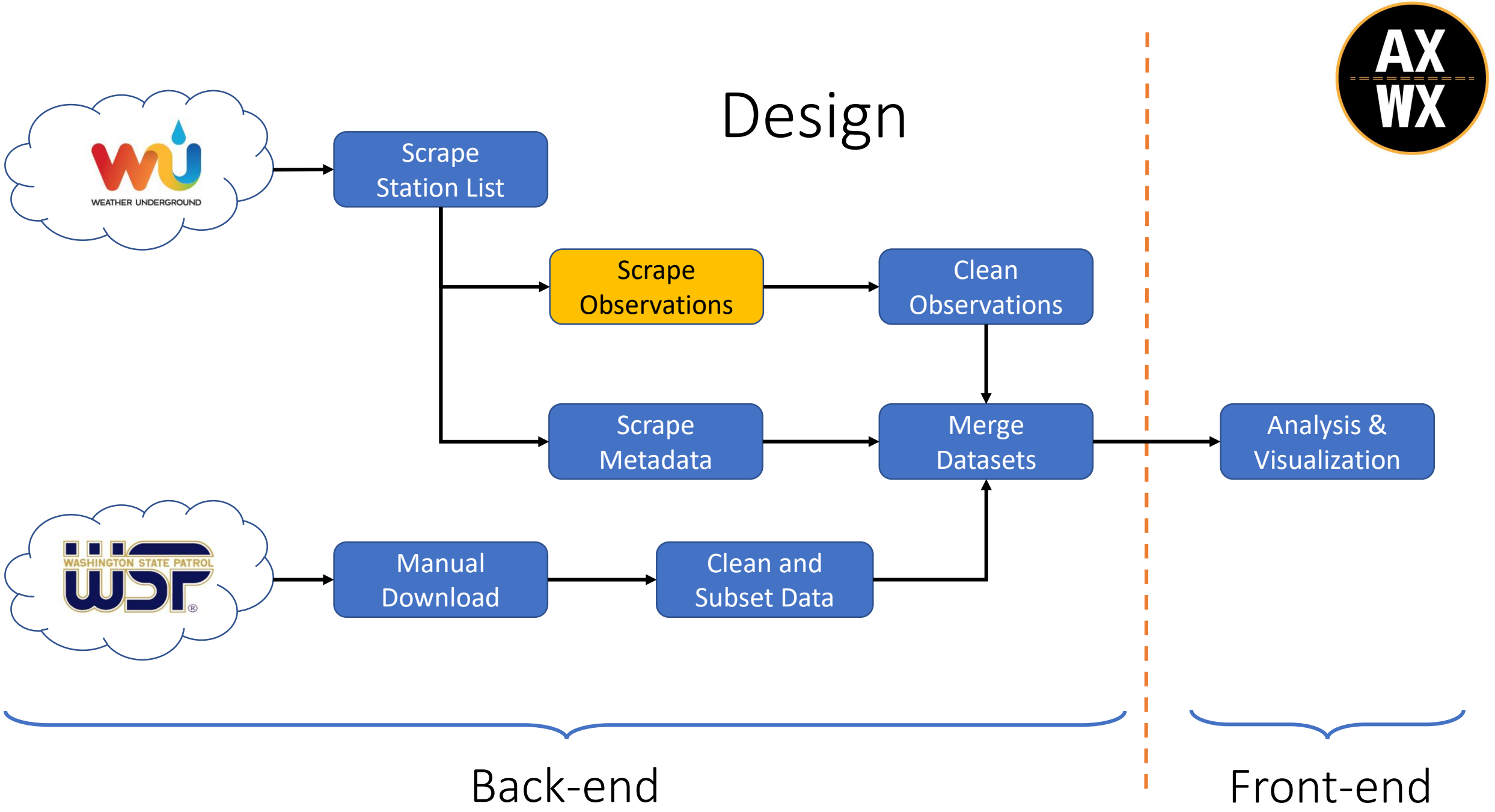
https://api.wunderground.com/weatherstation/WXDailyHistory.asp?ID=KWASEATT1735&forma

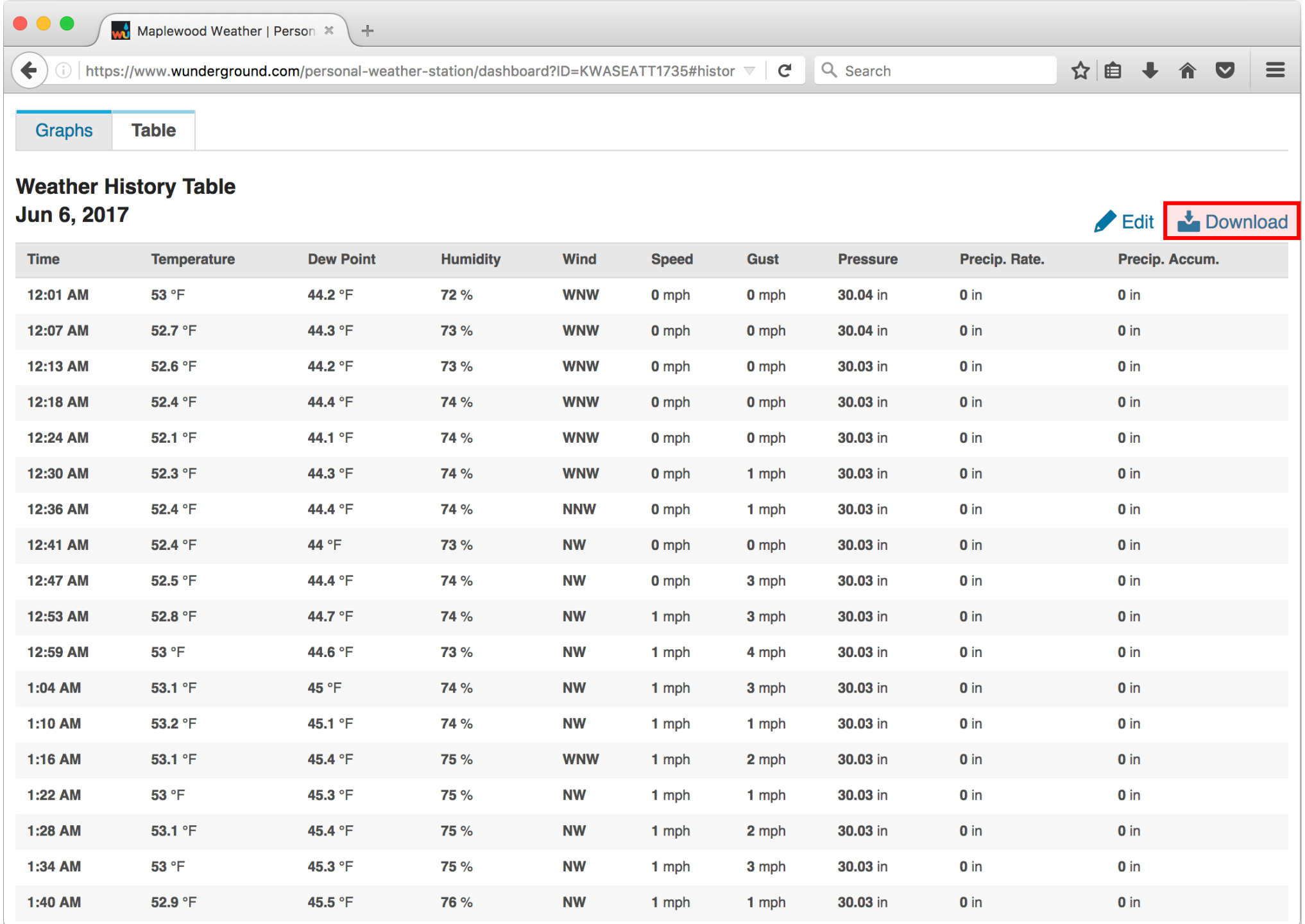
Search

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
– <day_observations>
  – <current_observation>
      <credit>Weather Underground Personal Weather Station</credit>
      <credit_URL>http://wunderground.com/weatherstation/</credit_URL>
    – <image>
      – <url>
          http://icons.wunderground.com/graphics/bh-wui_logo.gif
        </url>
        <title>Weather Underground</title>
        <link>http://wunderground.com/weatherstation/</link>
      </image>
    – <location>
        <full> UW ERIS Weather Station #2, Seattle, WA, WA</full>
        <neighborhood> UW ERIS Weather Station #2</neighborhood>
        <city>Seattle</city>
        <state>WA</state>
        <zip>98195</zip>
        <latitude>47.64879227</latitude>
        <longitude>-122.31082153</longitude>
        <elevation>30 ft</elevation>
      </location>
      <station_id>KWASEATT1735</station_id>
      <station_type/>
      <observation_time>Last Updated on June 6, 12:00 AM PDT</observation_time>
      <observation_time_rfc822>Tue, 06 Jun 2017 07:00:00 GMT</observation_time_rfc822>
      <weather/>
      <temperature_string>60.9 F ( C)</temperature_string>
      <temp_f>60.9</temp_f>
      <temp_c>16.1</temp_c>
      <relative_humidity>49</relative_humidity>
      <wind_string>From the ENE at 1.0 Gusting to 3.0 MPH</wind_string>
      <wind_dir>ENE</wind_dir>
      <wind_degrees>64</wind_degrees>
```

Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end



Graphs | Table

Weather History Table
Jun 6, 2017

Edit | Download

Time	Temperature	Dew Point	Humidity	Wind	Speed	Gust	Pressure	Precip. Rate.	Precip. Accum.
12:01 AM	53 °F	44.2 °F	72 %	WNW	0 mph	0 mph	30.04 in	0 in	0 in
12:07 AM	52.7 °F	44.3 °F	73 %	WNW	0 mph	0 mph	30.04 in	0 in	0 in
12:13 AM	52.6 °F	44.2 °F	73 %	WNW	0 mph	0 mph	30.03 in	0 in	0 in
12:18 AM	52.4 °F	44.4 °F	74 %	WNW	0 mph	0 mph	30.03 in	0 in	0 in
12:24 AM	52.1 °F	44.1 °F	74 %	WNW	0 mph	0 mph	30.03 in	0 in	0 in
12:30 AM	52.3 °F	44.3 °F	74 %	WNW	0 mph	1 mph	30.03 in	0 in	0 in
12:36 AM	52.4 °F	44.4 °F	74 %	NNW	0 mph	1 mph	30.03 in	0 in	0 in
12:41 AM	52.4 °F	44 °F	73 %	NW	0 mph	0 mph	30.03 in	0 in	0 in
12:47 AM	52.5 °F	44.4 °F	74 %	NW	0 mph	3 mph	30.03 in	0 in	0 in
12:53 AM	52.8 °F	44.7 °F	74 %	NW	1 mph	3 mph	30.03 in	0 in	0 in
12:59 AM	53 °F	44.6 °F	73 %	NW	1 mph	4 mph	30.03 in	0 in	0 in
1:04 AM	53.1 °F	45 °F	74 %	NW	1 mph	3 mph	30.03 in	0 in	0 in
1:10 AM	53.2 °F	45.1 °F	74 %	NW	1 mph	1 mph	30.03 in	0 in	0 in
1:16 AM	53.1 °F	45.4 °F	75 %	WNW	1 mph	2 mph	30.03 in	0 in	0 in
1:22 AM	53 °F	45.3 °F	75 %	NW	1 mph	1 mph	30.03 in	0 in	0 in
1:28 AM	53.1 °F	45.4 °F	75 %	NW	1 mph	2 mph	30.03 in	0 in	0 in
1:34 AM	53 °F	45.3 °F	75 %	NW	1 mph	3 mph	30.03 in	0 in	0 in
1:40 AM	52.9 °F	45.5 °F	76 %	NW	1 mph	1 mph	30.03 in	0 in	0 in

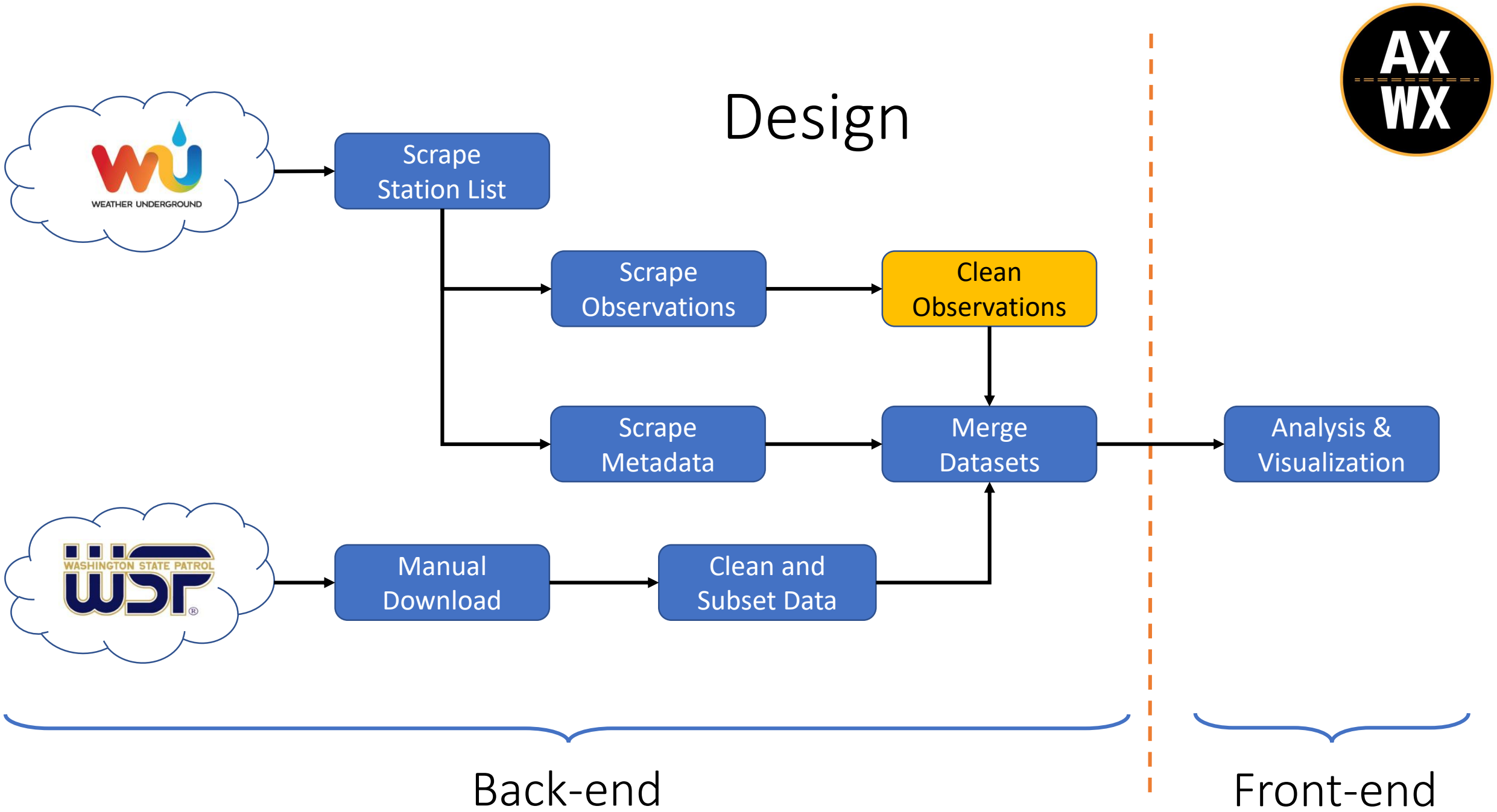
https://www.wunderground.c

https://www.wunderground.com/weatherstation/WXDailyHistory.asp?ID=KWASEATT1735&day

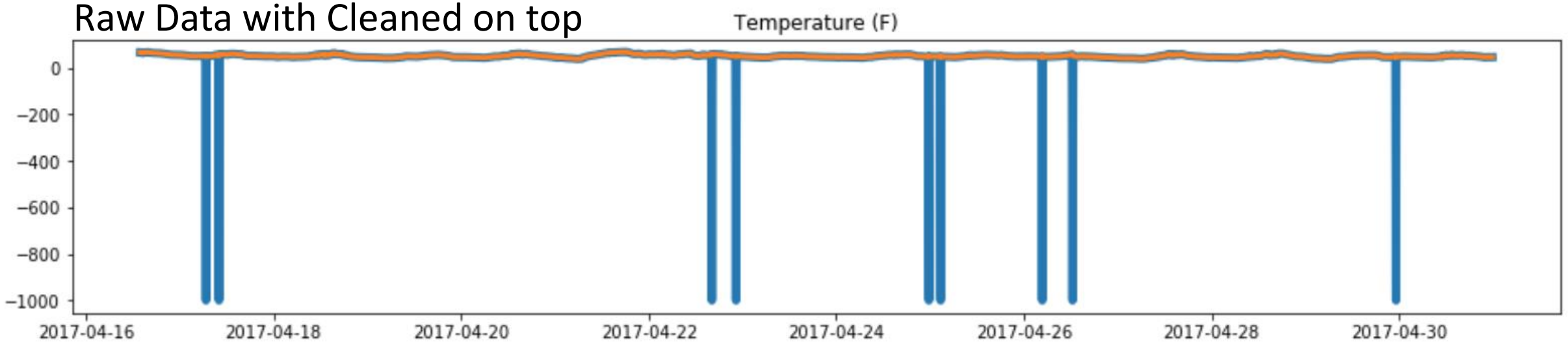
Search

Time,TemperatureF,DewpointF,PressureIn,WindDirection,WindDirectionDegrees,WindSpeedMPH,WindSpeedGustMPH,Humidity,HourlyPrecipIn,Conditions,Clouds,daily
2017-06-06 00:00:00,60.9,41.6,30.06,ENE,64,1.0,3.0,49,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:00:00,
2017-06-06 00:05:00,60.9,42.1,30.06,ESE,118,1.0,3.0,50,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:05:00,
2017-06-06 00:15:00,60.6,41.8,30.06,North,355,1.0,1.0,50,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:15:00,
2017-06-06 00:20:00,60.4,41.7,30.06,NNW,342,2.0,2.0,50,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:20:00,
2017-06-06 00:30:00,60.3,42.6,30.05,NNE,19,1.0,1.0,52,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:30:00,
2017-06-06 00:35:00,60.2,43.5,30.06,East,84,1.0,1.0,54,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:35:00,
2017-06-06 00:45:00,60.1,43.4,30.06,East,89,2.0,2.0,54,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:45:00,
2017-06-06 00:50:00,60.3,43.1,30.06,SSW,209,1.0,2.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 07:50:00,
2017-06-06 01:00:00,60.0,42.8,30.06,WSW,250,3.0,3.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:00:00,
2017-06-06 01:05:00,59.9,42.7,30.06,West,269,2.0,3.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:05:00,
2017-06-06 01:15:00,59.8,42.6,30.05,West,269,1.0,1.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:15:00,
2017-06-06 01:20:00,60.0,41.8,30.05,West,269,1.0,1.0,51,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:20:00,
2017-06-06 01:30:00,60.1,43.4,30.05,West,274,0.0,0.0,54,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:30:00,
2017-06-06 01:35:00,60.0,42.8,30.05,NW,321,0.0,2.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:35:00,
2017-06-06 01:45:00,60.2,43.0,30.06,SW,215,1.0,1.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:45:00,
2017-06-06 01:50:00,59.9,42.7,30.05,East,80,0.0,0.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:50:00,
2017-06-06 01:55:00,59.5,42.3,30.05,NNW,341,1.0,1.0,53,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 08:55:00,
2017-06-06 02:05:00,59.1,42.9,30.06,NNW,346,1.0,1.0,55,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:05:00,
2017-06-06 02:10:00,59.0,43.3,30.05,North,355,2.0,2.0,56,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:10:00,
2017-06-06 02:20:00,58.6,43.9,30.05,ENE,72,1.0,1.0,58,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:20:00,
2017-06-06 02:25:00,58.5,43.3,30.05,NNE,19,2.0,2.0,57,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:25:00,
2017-06-06 02:35:00,58.6,45.2,30.05,ENE,63,1.0,2.0,61,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:35:00,
2017-06-06 02:50:00,58.4,44.1,30.05,NNW,342,2.0,4.0,59,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:50:00,
2017-06-06 02:55:00,58.2,44.4,30.05,ENE,69,2.0,4.0,60,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 09:55:00,
2017-06-06 03:05:00,58.0,44.6,30.05,ENE,65,2.0,2.0,61,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:05:00,
2017-06-06 03:10:00,58.1,45.6,30.05,ESE,116,1.0,1.0,63,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:10:00,
2017-06-06 03:15:00,58.2,46.1,30.06,South,172,1.0,1.0,64,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:15:00,
2017-06-06 03:25:00,58.0,44.6,30.05,NNE,15,1.0,1.0,61,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:25:00,
2017-06-06 03:30:00,58.0,44.6,30.05,NE,55,2.0,2.0,61,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:30:00,
2017-06-06 03:40:00,57.6,45.5,30.04,East,82,2.0,2.0,64,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:40:00,
2017-06-06 03:45:00,57.7,45.6,30.04,East,82,2.0,2.0,64,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:45:00,
2017-06-06 03:55:00,57.6,46.3,30.04,East,100,2.0,3.0,66,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 10:55:00,
2017-06-06 04:00:00,57.6,47.1,30.04,South,169,2.0,2.0,68,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 11:00:00,
2017-06-06 04:10:00,57.2,47.5,30.04,East,100,1.0,1.0,70,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 11:10:00,
2017-06-06 04:15:00,57.1,45.8,30.04,North,355,1.0,2.0,66,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 11:15:00,
2017-06-06 04:25:00,57.2,44.7,30.03,NE,46,1.0,1.0,63,0.00,,,0.00,0,Wunderground v.1.15,2017-06-06 11:25:00,

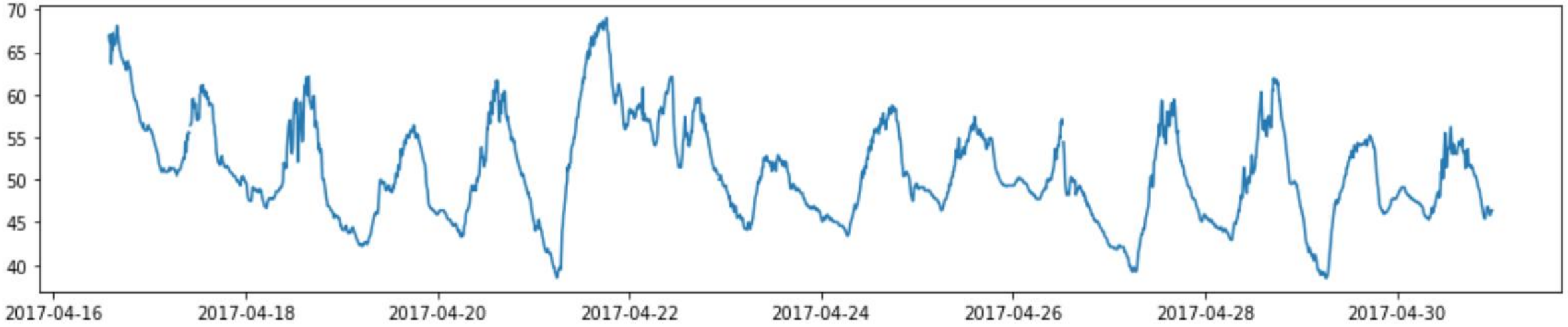
Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

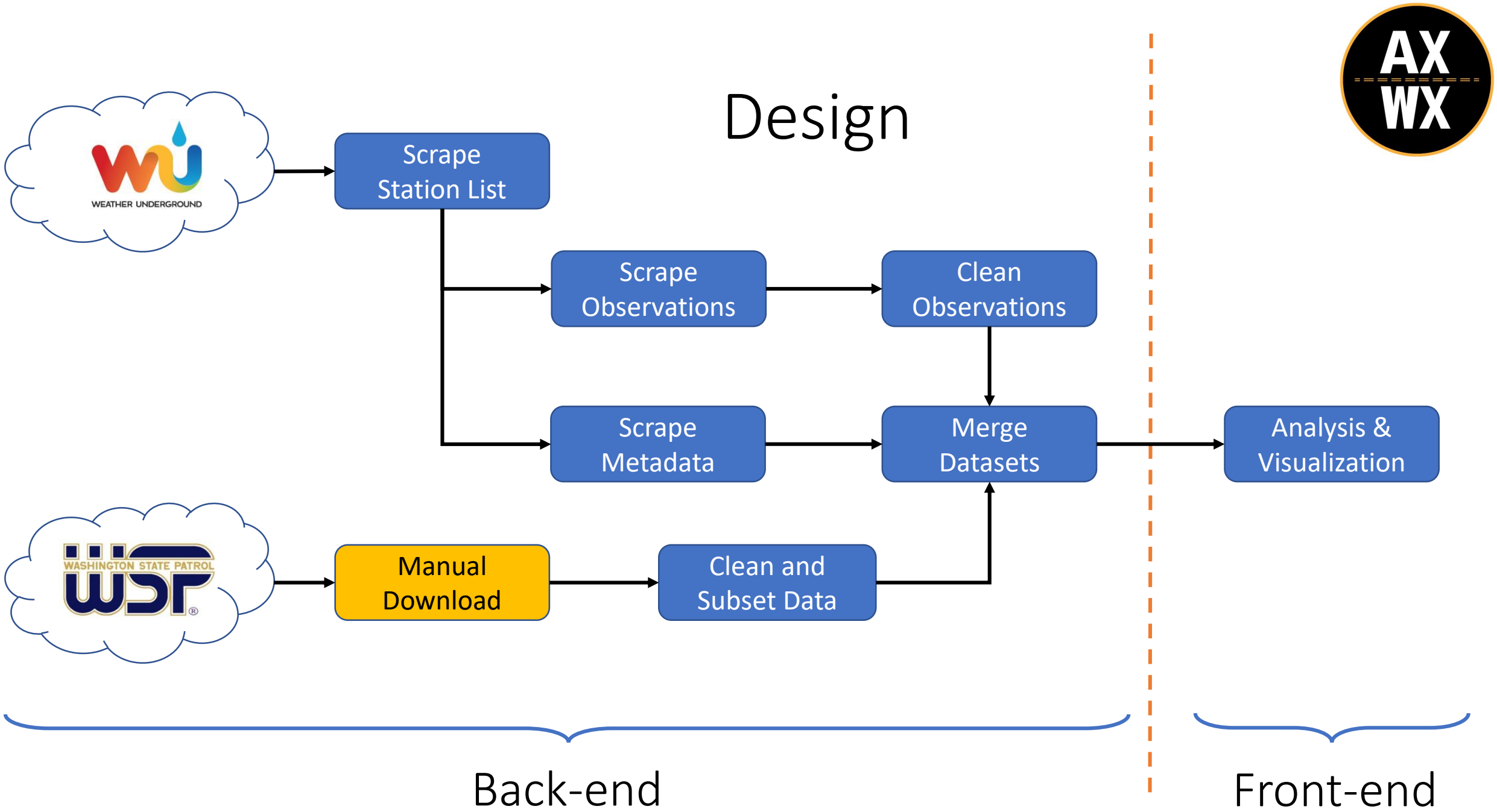
Raw Data with Cleaned on top



Cleaned Data

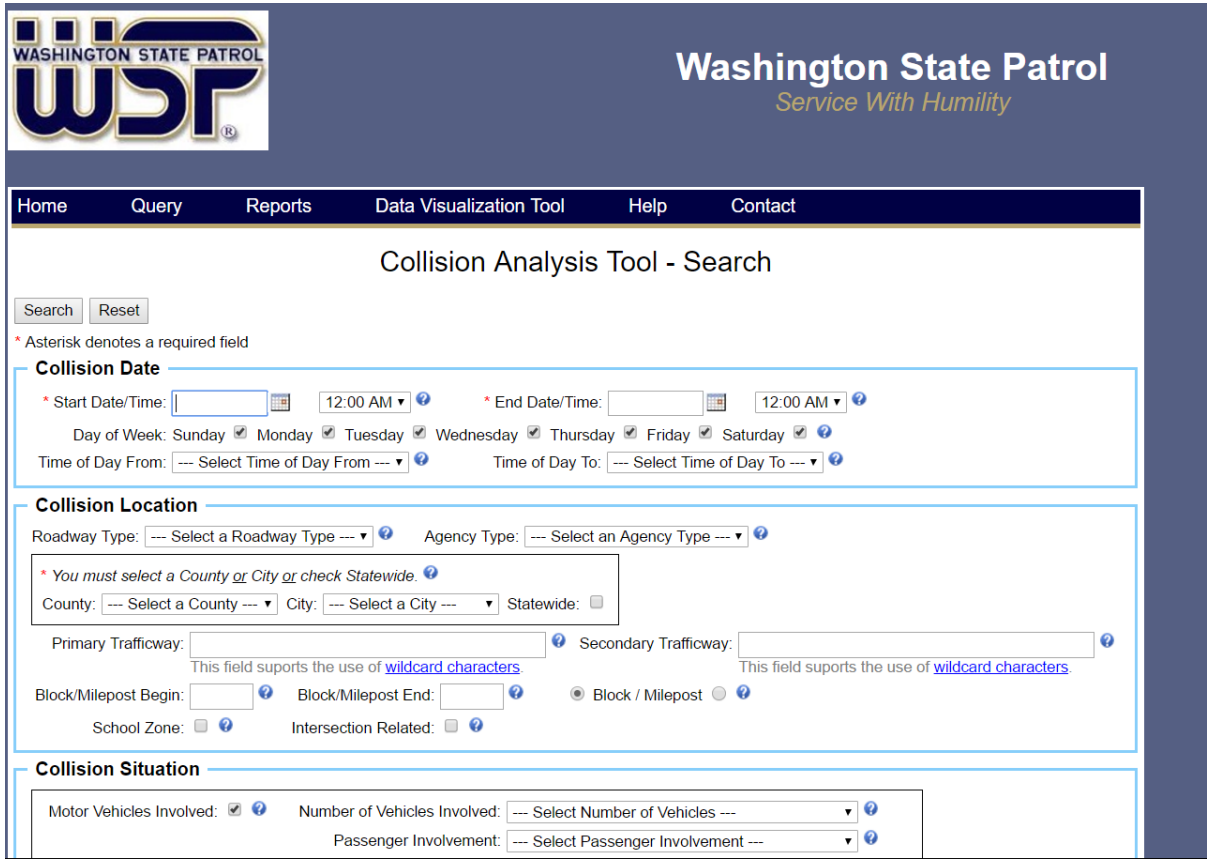


Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

WSP Data Processing – Obtaining raw data

- Online querying tool pulls data from 2013 onwards
  - Requests prior to 2013 made through email

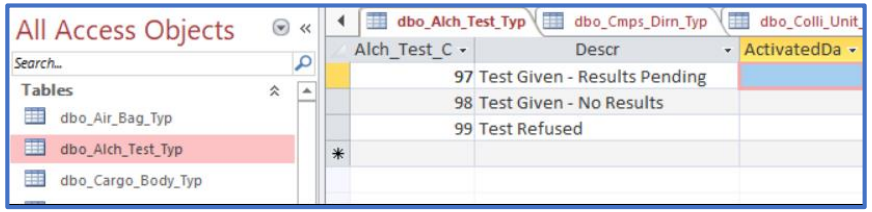


WSP Data Processing

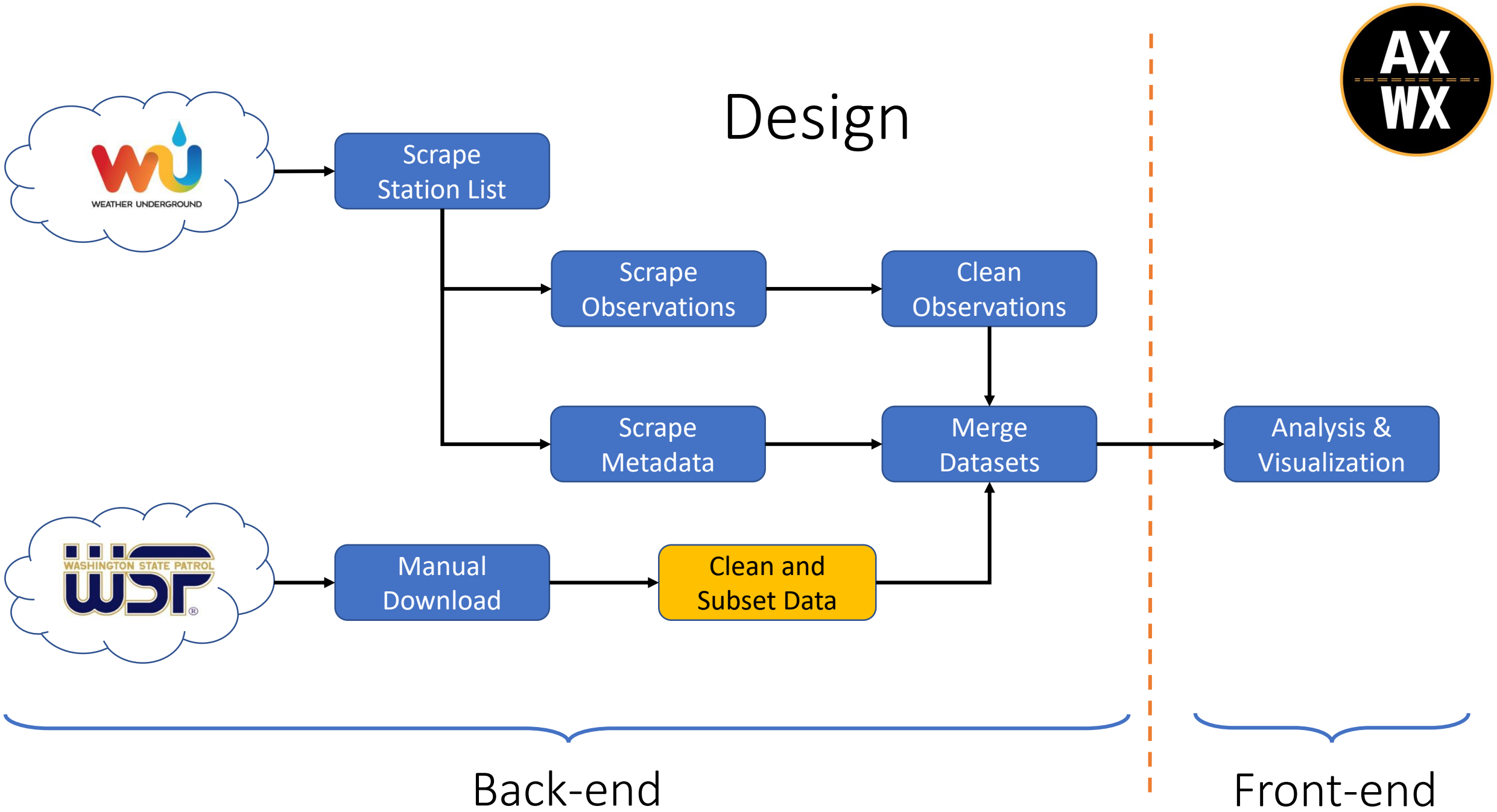
- Raw WSP data in CSV form
- Data Dictionary and Look-Up table included for decoding

	A	B	C	D	E	F	G	H	I	J	K	L
1	Colli_Rpt_	Colli_Rpt_	Colli_Rpt_Num_Trans_Date_Time	Colli_Rpt_	Colli_Rpt_Num_Colli_Entry_Date_Time	Colli_Rpt_	Colli_Rpt_	Colli_Rpt_	Location_I	Location_(	Location_I	Location_I
2	E284926	1669246	11/14/2013 16:01	1	11/14/2013 16:01	4		1	63	E284926	SR 504	0.75
3	E284926	1669246	11/14/2013 16:01	1	11/14/2013 16:01	4		1	63	E284926	SR 504	0.75
4	E284927	1669247	11/14/2013 16:01	1	11/14/2013 16:01	4		1	64	E284927	STATE ROUTE 504	
5	E284928	1669248	11/14/2013 16:01	2	11/14/2013 16:01	4		1	65	E284928	Key Peninsula Hwy N	
6	E284929	1669250	11/14/2013 16:01	2	11/14/2013 16:01	4		1	67	E284929	12TH AVE	500
7	E284929	1669250	11/14/2013 16:01	2	11/14/2013 16:01	4		1	67	E284929	12TH AVE	500
8	E284931	1669252	11/14/2013 16:02	3	11/14/2013 16:02	4		1	69	E284931	228 ST SE	
9	E284931	1669252	11/14/2013 16:02	3	11/14/2013 16:02	4		1	69	E284931	228 ST SE	
10	E284934	1669255	11/14/2013 16:02	1	11/14/2013 16:02	4		1	72	E284934	SR 512	0.5
11	E284934	1669255	11/14/2013 16:02	1	11/14/2013 16:02	4		1	72	E284934	SR 512	0.5
12	E284934	1669255	11/14/2013 16:02	1	11/14/2013 16:02	4		1	72	E284934	SR 512	0.5
13	E284935	1669256	11/14/2013 16:02	1	11/14/2013 16:02	4		1	73	E284935	SR 512	0.1
14	E284935	1669256	11/14/2013 16:02	1	11/14/2013 16:02	4		1	73	E284935	SR 512	0.1
15	E284935	1669256	11/14/2013 16:02	1	11/14/2013 16:02	4		1	73	E284935	SR 512	0.1
16	E284938	1669259	11/14/2013 16:02	3	11/14/2013 16:02	4		1	76	E284938	S GRADY WAY	
17	E284938	1669259	11/14/2013 16:02	3	11/14/2013 16:02	4		1	76	E284938	S GRADY WAY	
18	E284938	1669259	11/14/2013 16:02	3	11/14/2013 16:02	4		1	76	E284938	S GRADY WAY	
19	E284940	1669261	11/14/2013 16:02	1	11/14/2013 16:02	4		1	78	E284940	HWY 167	
20	E284940	1669261	11/14/2013 16:02	1	11/14/2013 16:02	4		1	78	E284940	HWY 167	
21	E284940	1669261	11/14/2013 16:02	1	11/14/2013 16:02	4		1	78	E284940	HWY 167	
22	E284940	1669261	11/14/2013 16:02	1	11/14/2013 16:02	4		1	78	E284940	HWY 167	
23	E284941	1669262	11/14/2013 16:03	3	11/14/2013 16:03	4		1	79	E284941	148 AVE SI	300
24	E284941	1669262	11/14/2013 16:03	3	11/14/2013 16:03	4		1	79	E284941	148 AVE SI	300
25	E284944	1669265	11/14/2013 16:03	3	11/14/2013 16:03	4		1	82	E284944	106TH AVE NE	

	Table	Column	Primary Key	Data Type	Length	Precision	Scale	Nullable	Description
2	**Colli_Rpt_Num**	Colli_Rpt_Num		CHAR	7			N	The collision report number
3		Colli_Surr_Key	Y	INT				N	System-generated unique integer value used as primary key for Colli_Rpt_Num table
3	Table that contains identifying information regarding a Police Traffic Collision Report form or Civilian Vehicle Collision Report								
4		Trans_Date_Time		DATETIME				N	The date/time the record was last updated in the database
5		Colli_Rpt_Typ_Surr_Key		TINYINT				N	Jurisdiction where the crash occurred. The values can be found in the Colli_Rpt_Typ lookup table
6		Colli_Entry_Date_Time		DATETIME				Y	The date/time the record was entered in the database
7		QA_Typ_Cd		TINYINT				Y	The current status of a collision report, DA (Data Analysis) or QA (Quality Assurance). The values can be found in the QA_Typ lookup table
7		QA_Date_Time		DATETIME				Y	The date/time a collision report was been saved within the CLAS application's Quality Assurance mode
8		Colli_Rpt_Src_Typ_Cd		TINYINT				Y	Crash report source. The values can be found in the Colli_Rpt_Src_Typ lookup table
9									
10	**Location**	Id	Y	INT				N	System-generated unique integer value used as primary key for Location table
11									
12	Detail table that contains information regarding the physical location of a collision	Colli_Rpt_Num		CHAR	7			N	The crash report number
13		Pri_Trfcway		CHAR	20			Y	The primary traffic way of the crash location
13		Dirn_From_Ref_Pt		DECIMAL		7	3	Y	Distance from the reference point where the crash occurred
14		Ref_Pt_Miles_Feet_Ind		CHAR	1			Y	Indicates if the distance from reference point is in miles of feet. Value will be either 'M' for miles of 'F' for
15		Cmps_Dirn_Typ_Cd		CHAR	1			Y	Compass direction of the crash location. The values can be found in the Cmps_Dirn_Typ lookup table
16		Secondary_Trfcway		CHAR	20			Y	The secondary traffic way of the crash location
17		Block_Num		CHAR	6			Y	The block number where the crash occurred if Ref_Pt_Block_MilePost_Ind value is 'B'
18		MilePost		DECIMAL		6	3	Y	Milepost where the crash occurred if Ref_Pt_Block_MilePost_Ind value is 'M'
19		Ref_Pt_Block_MilePost_Ind		CHAR	1			Y	Indicates if the reference point is a block number or milepost. Value will be 'B' for block or 'M' for MilePost
20		Colli_Rpt_Typ_Surr_Key		TINYINT				Y	Jurisdiction where the crash occurred. The values can be found in the Colli_Rpt_Typ lookup table
21									



Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

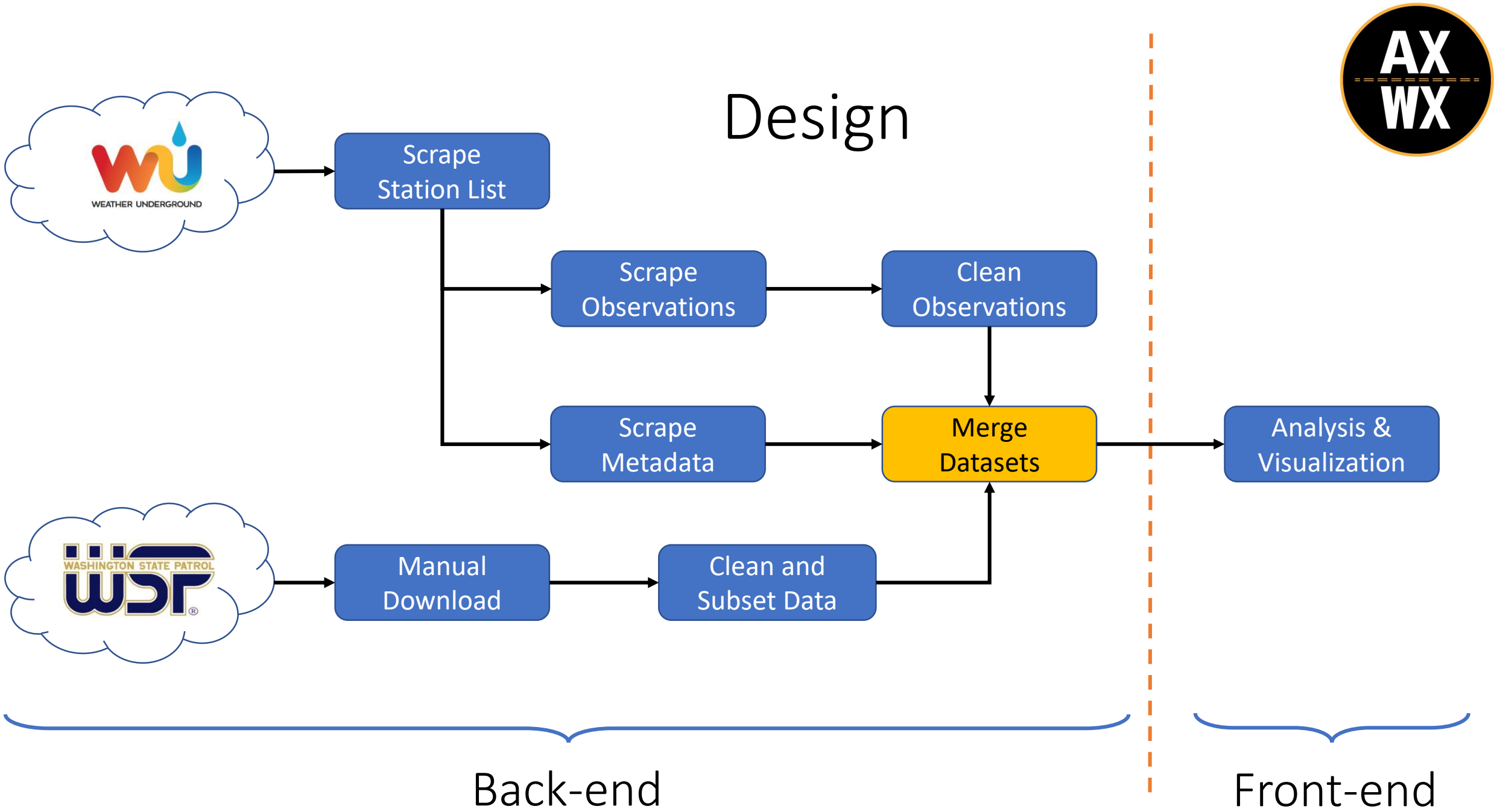
WSP Data Processing

- WSP data cleaned in Python by identifying relevant fields
- Using references to data dictionaries and lookup tables for relevant fields

Out[5]:

	Unnamed: 0	lat	lon	date	time_of_day	month	day_of_week	hour	driver_restraint_type	passenger_restraint_typ
0	0	47.674063	-122.202327	2016-05-01	12:13:00	5	6	12	Lap & Shoulder Used	NaN
1	1	47.674063	-122.202327	2016-05-01	12:13:00	5	6	12	Lap & Shoulder Used	NaN
2	2	47.534763	-122.295927	2016-05-01	17:13:00	5	6	17	Lap & Shoulder Used	NaN
3	3	47.534763	-122.295927	2016-05-01	17:13:00	5	6	17	Lap & Shoulder Used	NaN
4	4	47.534763	-122.295927	2016-05-01	17:13:00	5	6	17	Lap & Shoulder Used	NaN
5	5	47.534763	-122.295927	2016-05-01	17:13:00	5	6	17	NaN	NaN
6	6	47.431542	-122.295415	2016-05-01	07:41:00	5	6	7	Lap & Shoulder Used	NaN
7	7	47.431542	-122.295415	2016-05-01	07:41:00	5	6	7	NaN	NaN
8	8	47.578178	-122.209709	2016-05-01	01:47:00	5	6	1	No Restraints Used	NaN
9	9	47.578178	-122.209709	2016-05-01	01:47:00	5	6	1	Lap & Shoulder Used	NaN
10	10	47.602054	-122.325379	2016-05-01	17:37:00	5	6	17	Lap & Shoulder Used	Lap & Shoulder Used
11	11	47.602054	-122.325379	2016-05-01	17:37:00	5	6	17	Lap & Shoulder Used	Lap & Shoulder Used
12	12	47.602054	-122.325379	2016-05-01	17:37:00	5	6	17	Lap & Shoulder Used	Lap & Shoulder Used
13	13	47.602054	-122.325379	2016-05-01	17:37:00	5	6	17	Lap & Shoulder Used	Lap & Shoulder Used
14	14	47.602054	-122.325379	2016-05-01	17:37:00	5	6	17	Lap & Shoulder Used	Lap & Shoulder Used
15	15	47.659482	-122.204806	2016-05-01	16:07:00	5	6	16	Unknown	NaN
16	16	47.659482	-122.204806	2016-05-01	16:07:00	5	6	16	Lap & Shoulder Used	Lap & Shoulder Used

Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

Merging the data

- WU and WSP datasets merged by identifying closest weather station to coordinates of accident occurrence within 3 miles

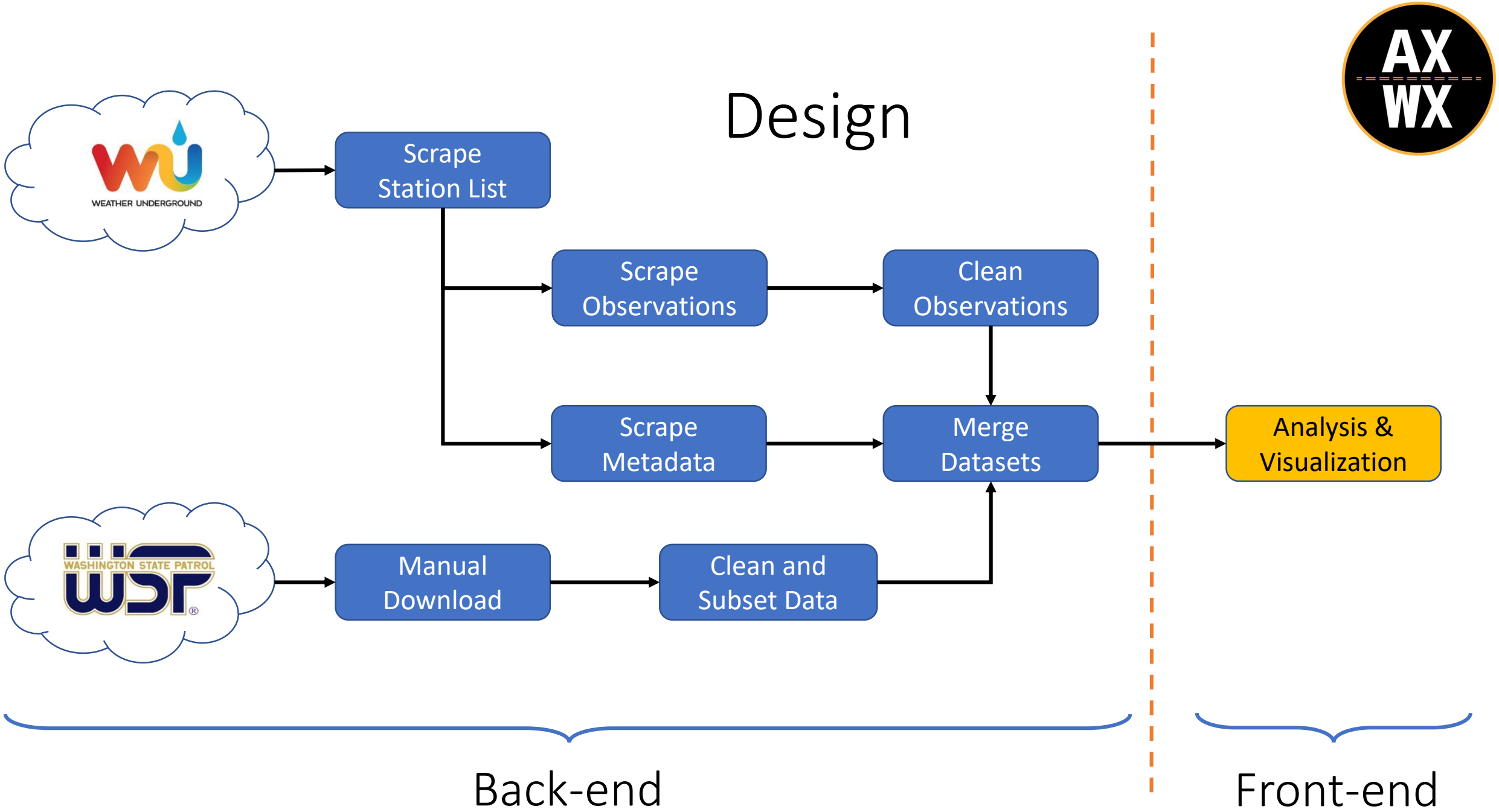
```
In [115]: raw_data.columns

Out[115]: Index(['airbag', 'alcohol_test_given', 'contributing_factor_1',
       'contributing_factor_2', 'contributing_factor_3', 'current_weather',
       'cyclist_injury', 'cyclist_present', 'date', 'day_of_week',
       'driver_injury', 'driver_restraint_type', 'hour', 'lat',
       'lighting_conditions', 'lon', 'month', 'passenger_injury',
       'passenger_restraint_type', 'pedestrian_injury', 'pedestrian_present',
       'posted_speed_limit', 'roadway_characterization',
       'roadway_surface_condition', 'roadway_surface_type', 'sobriety_type',
       'time_of_day', 'vehicle_action', 'wx_DewpointF_last_1hr_avg',
       'wx_DewpointF_last_1hr_avg_decrease',
       'wx_DewpointF_last_1hr_avg_increase', 'wx_DewpointF_last_1hr_change',
       'wx_DewpointF_latest', 'wx_Humidity_last_1hr_avg',
       'wx_Humidity_last_1hr_avg_decrease',
       'wx_Humidity_last_1hr_avg_increase', 'wx_Humidity_last_1hr_change',
       'wx_Humidity_latest', 'wx_PrecipRate_inhr_last_1hr',
       'wx_PrecipRate_inhr_latest_max', 'wx_PressureIn_last_1hr_avg',
       'wx_PressureIn_last_1hr_change', 'wx_PressureIn_latest',
       'wx_TemperatureF_last_1hr_avg', 'wx_TemperatureF_last_1hr_avg_decrease',
       'wx_TemperatureF_last_1hr_avg_increase',
       'wx_TemperatureF_last_1hr_change', 'wx_TemperatureF_latest',
       'wx_WindSpeedGustMPH_last_1hr_max', 'wx_WindSpeedGustMPH_latest',
       'wx_WindSpeedMPH_last_1hr_avg', 'wx_WindSpeedMPH_latest',
       'wx_mean_station_dist_mi', 'wx_station_count', 'wx_unique_event_id',
       'rain_bin', 'severe_ax', 'accident_id', 'temp_round'],
      dtype='object')
```

```
In [114]: raw_data
```

Out[114]:

	airbag	alcohol_test_given	contributing_factor_1	contributing_factor_2	contributing_factor_3	current_weather	cyclist_injury	cyclist_pre
0	Not Deployed	NaN	Did Not Grant RW to Vehicle	NaN	NaN	Clear or Partly Cloudy	NaN	NaN
1	Not Deployed	NaN	None	NaN	NaN	Clear or Partly Cloudy	NaN	NaN
2	Front Airbag Deployed	NaN	Improper Passing	NaN	NaN	Clear or Partly Cloudy	NaN	NaN
3	Not Deployed	NaN	None	NaN	NaN	Clear or Partly Cloudy	NaN	NaN
4	Not Deployed	NaN	None	NaN	NaN	Clear or Partly Cloudy	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	Clear or Partly Cloudy	NaN	NaN

Design


WEATHER UNDERGROUND
Scrape Station List
Scrape Observations
Clean Observations
Scrape Metadata
Merge Datasets
Analysis & Visualization
WASHINGTON STATE PATROL
Manual Download
Clean and Subset Data
Back-end
Front-end

Project Structure

https://github.com/rexthompson/axwx

Lessons Learned and Future Work

- Lessons learned
  - Valuable hands-on experience with Git and package structure
  - Respect the time it takes to acquire/clean/merge data
  - Technology review is critical
  - Web scraping is fun!

- Future work
  - Continue with exploratory data analysis
  - Further package developments (WSP geolocation, WU quality control)
  - Other WU PWS data applications (e.g. WA DNR prescribed burns)