

Caltrans PeMS Workshop

presented to
Caltrans District 5

presented by
Aldo Tudela

atudelarivadenevra@camsys.com



October 30, 2018

Think  Forward

Workshop Objectives

- Provide guidance such that attendees feel comfortable using the PeMS platform and its data.
- Describe general usage, tips, and identify resources and materials for future consultation.
- Give attendees real examples to use for the development of a corridor analysis study.

Content

- Introduction to PeMS
 - » Introduction
 - » How Does PeMS Work?
 - » Data Available
 - » System Overview
 - » Help & References

- Exercise 1 – Familiarizing with Dashboard

- Exercise 2 – Evaluate Data Quality on Corridor

- Exercise 3 - Evaluate Traffic Counts on Corridor

- Exercise 4 – Evaluate Typical Speeds on Corridor

- Exercise 5 – Downloading Raw Data

INTRODUCTION TO PEMS

Introduction

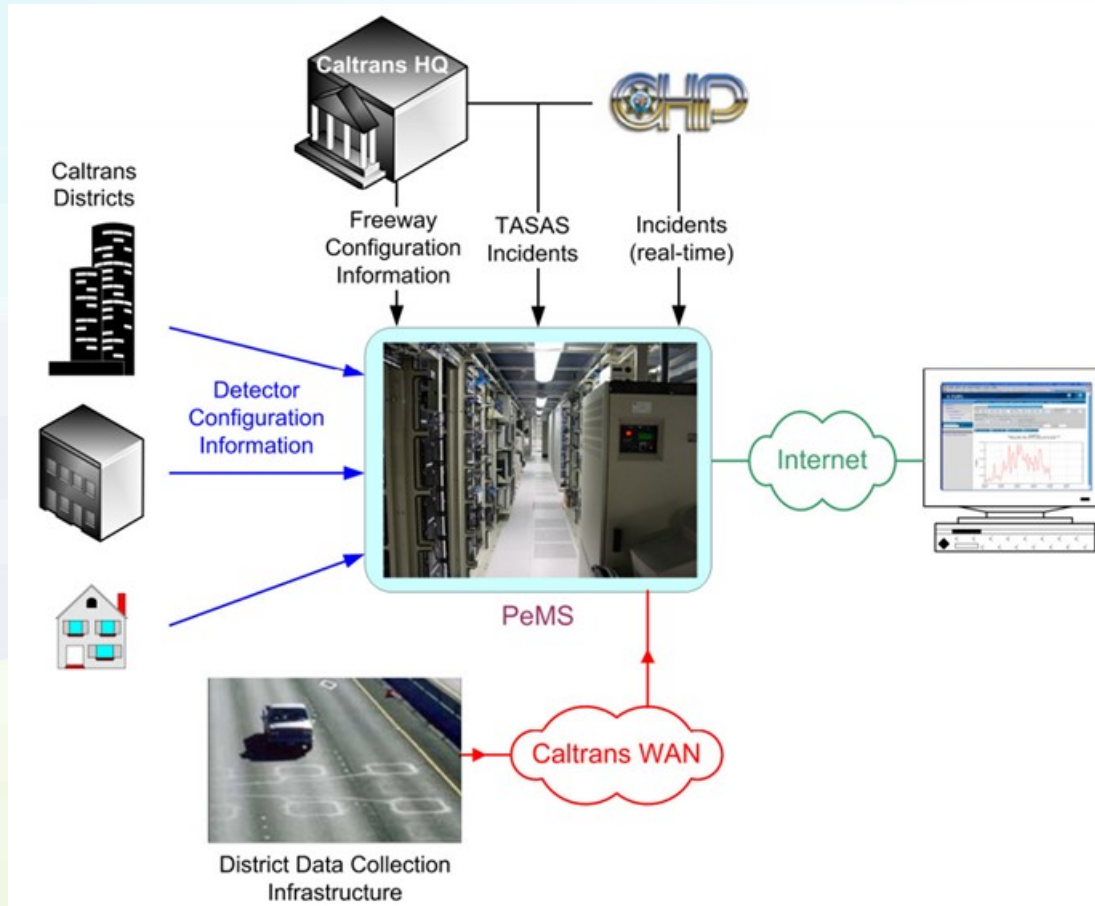
➤ What is PeMS?

- » PeMS stands for Performance Measurement System, it is Caltrans depository for real-time traffic data.
- » PeMS started in 1999 as a UC PATH research project.
- » Processes data of over 35,000 detectors, every 30 seconds.
- » Data is stored from inception (detector added) and never deleted, it has now over 12 Tb of data.

➤ How can I access it?

- » Go to <http://pems.dot.ca.gov/>
- » Create User Profile (User Name and Password).

How Does PeMS Work?



Data Sources

- Intelligent Transportation System (ITS) Vehicle Detector Stations (VDS)

- Traffic Counters
 - » Traffic Census Stations
 - » Weight-In-Motion (WIM) Sensors

- Other Data Sets
 - » California Highway Patrol (CHP) Incident data
 - » The Caltrans Traffic Accident Surveillance and Analysis System (TASAS) accident data (for Caltrans users only)
 - » Lane Closure information from the Caltrans Lane Closure System
 - » Electronic Toll Collection (ETC) Reader data (Bay Area only)
 - » Changeable Message Signs (real-time information only)
 - » Arterial Detector data and Timing Plans (limited data in District 11)
 - » Transit data such as routes and schedules, Automated Vehicle Location (AVL) and Automated Passenger Count (APC) data (limited data in District 11)

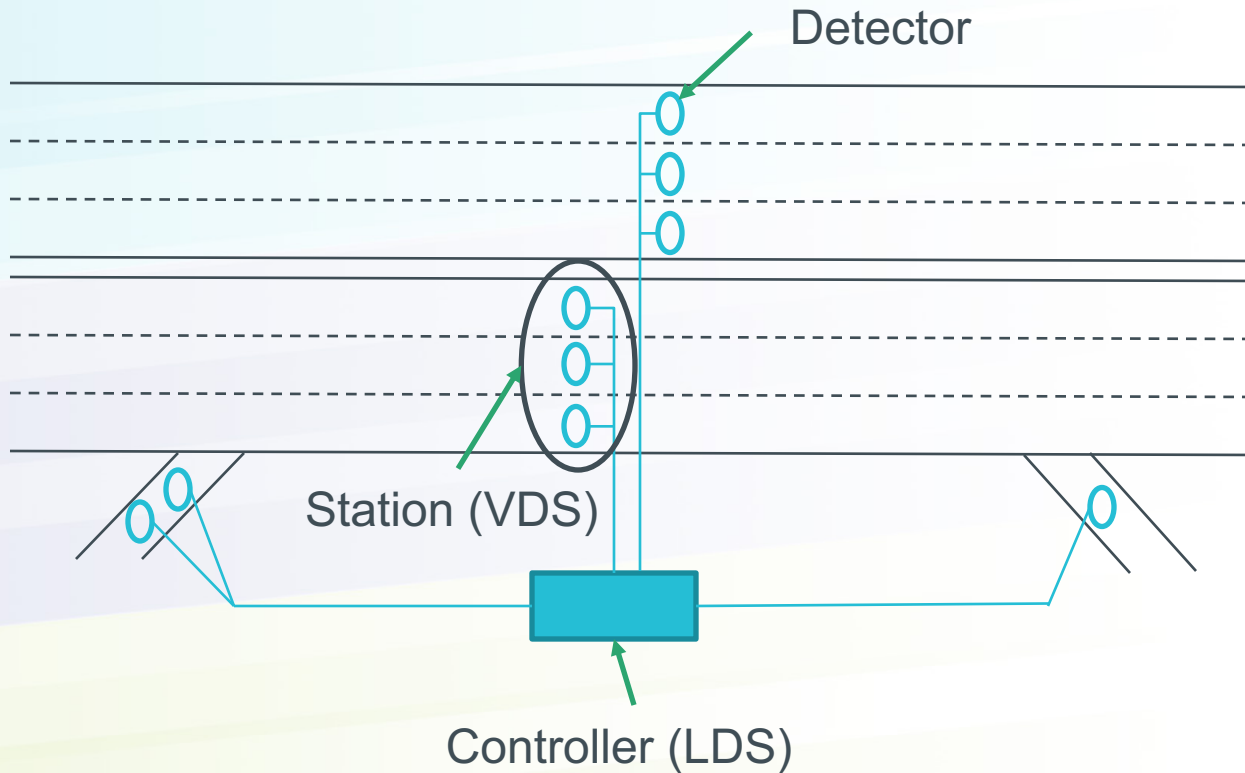
Data Available

- Computes standard transportation performance measures, such as **VMT**, **VHT**, **Delay** (expressed in vehicle-hours), and **Level of Service (LOS)**
- Calculates **travel time** and travel time reliability measures, such as the **Buffer Time Index**, **Travel Time Index**, and other descriptive statistics
- Produces **summary reports**, such as locations with low traffic flow or high VMT over several years
- **Imputes data** for missing or bad detector data in real-time.
- Provides **speed** as reported by detectors, or computed speed based on flow and occupancy if the detector does not report speed.

System Overview (1)

- The primary data source is the **vehicle detector stations** (VDS).
- PeMS compiles 30 second data of vehicle flow and occupancy from VDS.
- VDS data might have gaps (missing data sets, errors, etc). PeMS conducts diagnostic tests to determine quality of data.
- PeMS calculates missing data using data imputation methods.
- Once 30-second dataset is complete, PeMS aggregates these into 5-minute datapoints.
- PeMS uses 5-minute data to calculate performance measures and saves it on servers for user access.

System Overview (2)



Help & Resources

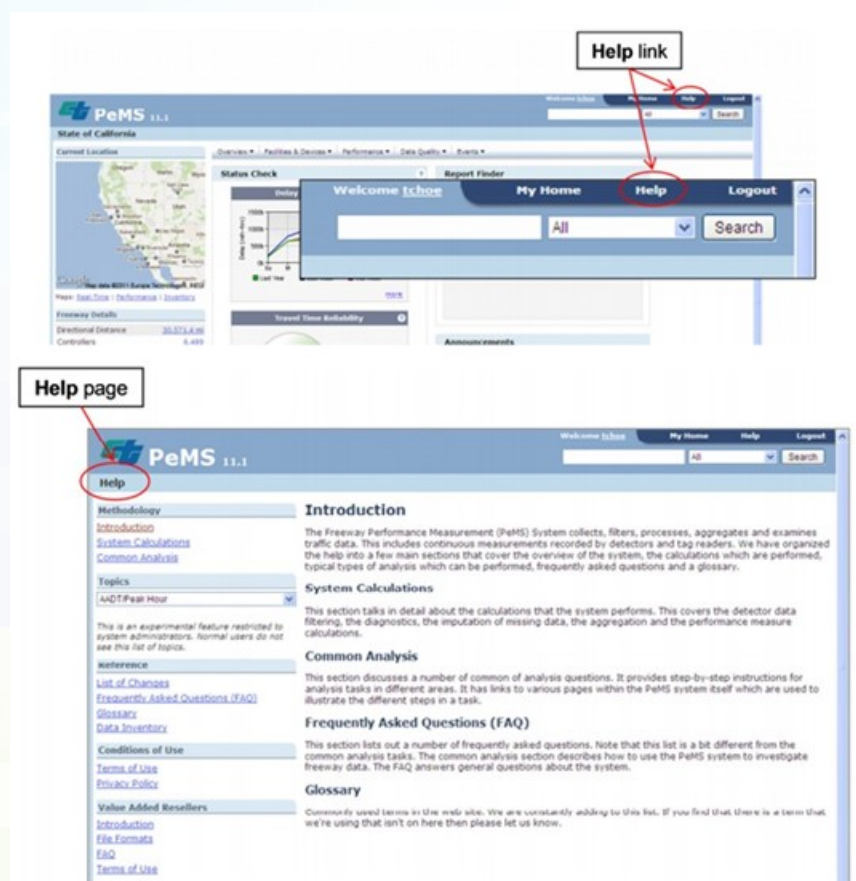
➤ PeMS User Manual:

http://pems.dot.ca.gov/PeMS_Intro_User_Guide_v5.pdf

➤ PeMS Help Page →

➤ PeMS FAQ:

http://pems.dot.ca.gov/?directory=Help&node=Help&content=var_faq



EXERCISE 1: FAMILIARIZING WITH DASHBOARD.



Dashboard

Reports Tab

Account Settings

Help Sources

Map view

PeMS 18.0
State of California

Current Location: Overview | Facilities & Devices | Performance | Data Quality | Events

Map view: Maps RealTime | Performance | Inventory

Freeway Details:
Directional Distance: 30,602.1 mi
Controllers: 5,965
Stations: 18,285
Detectors: 45,088
Traffic Census Stations: 16,527

Quick Links:
Jump to default page for district...
Jump to default page for county...
Jump to default page for city...
Jump to default page for freeway...

Featured Sections:
[Mobility Performance Report](#)
[Detector Health](#)
[CHP Incidents](#)
[Lane Closure System](#)
[Corridors](#)

Tools:
[Holidays](#)
[Data Clearinghouse](#)
[PeMS User Manual](#)

Status Check:

- Delay by Day of Week: Line chart showing delay (min-hr) by day of week for Last Year, Last Week, and This Week.
- Travel Time Reliability: Two gauge charts showing reliability for 5-10 AM and 3-8 PM.
- Detector Health: Line chart showing % Working vs. Statewide Goal (90%) and Yesterday's Health (67%).

Report Finder:

- Freeways: Select a freeway, Direction, Restrict Location - Optional, Select a report.
- Devices: (Empty section)

Announcements:

- Database Improvements** (September 12, 2018): All districts except D7 have been moved to the new higher-capacity Raw Collection Database.
- Database Maintenance** (September 04, 2018): One of PeMS' databases is being upgraded between Tuesday 09/04/18 and Friday 09/07/18.
- PeMS 18.0** (August 03, 2018): The 18.0 release contains 35 new features, improvements and fixes.
- Browser Security Info** (July 31, 2018): Chrome and other browsers are now displaying messages when users login to non-encrypted web sites like PeMS that use HTTP instead of HTTPS.



Finding data

Corridor analyzed

Map Viewer

Analysis Type

Performance > Spatial Analysis > 2-D

Configurable Parameters

Report type

Other data

Data Viewer

The screenshot displays the PeMS web application interface. At the top, the breadcrumb navigation path is "Performance > Spatial Analysis > 2-D". The main content area is divided into several sections:

- Map Viewer:** A map of California showing the location of Freeway US101-N in District 5, highlighted in blue.
- Freeway Details:** A table listing various attributes of the freeway, such as Directional Distance (266.2 mi), Stations (104), and Detectors (201).
- Performance Report:** A section titled "Speed for US101-N (66% Observed)" with a line chart showing speed in mph versus Postmile (Abs) for three lanes (Lane 1, Lane 2, Lane 3). The chart shows a significant drop in speed around postmile 100.
- Configurable Parameters:** A section on the right side of the report, including "VDS Type" (Mainline selected) and "Quantity" (Speed selected).
- Report type:** A row of buttons for exporting the report to various formats: DRAW PLOT, VIEW TABLE, EXPORT TEXT, EXPORT to XLS, EXPORT to PDF, and EXPORT to HTML.
- Other data:** A status bar at the bottom right of the report section showing 0 warnings, 13 alerts, and 0 errors.

EXERCISE 2: EVALUATE DATA QUALITY ON CORRIDOR

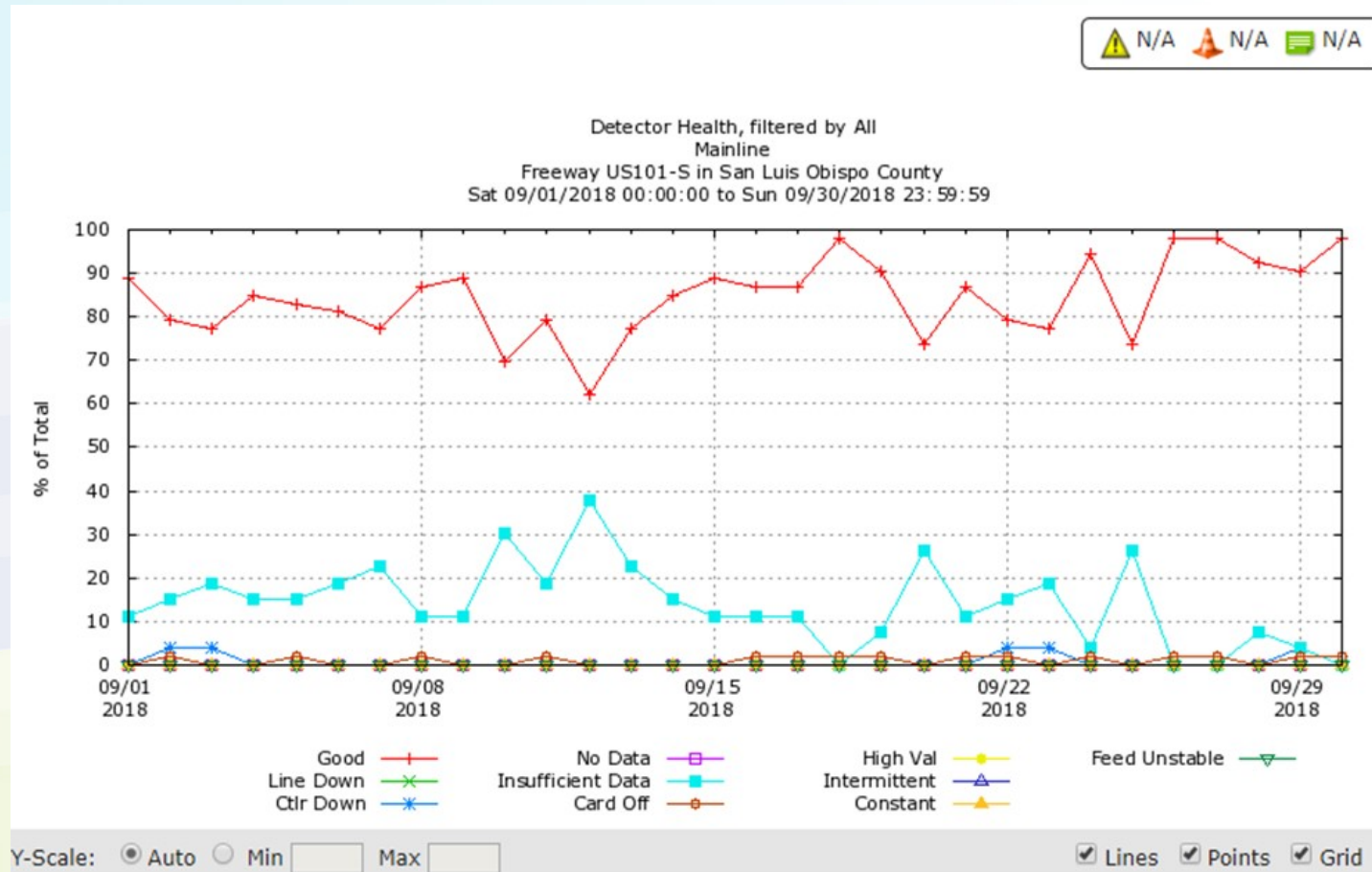
Objectives

- We want to assess the data quality for the US101 corridor in the San Luis Obispo County.
- We are interested in the Southbound direction only.
- We are interested in seeing the data quality of the corridor in the month of September.
- We want to assess the mainline and ramps separately .

Process

1. Select the corridor and geographic boundary of interest
2. Select data only for San Luis Obispo County
3. Navigate to Data Quality → Time Series
4. Modify dates for data

Output



Let's try new queries!

- Can we develop the same graph for on and off ramps?
- Can we take a look at specific data by day?
- What types of reports can be developed?

10 MINUTE BREAK!



EXERCISE 3 - EVALUATE TRAFFIC COUNTS ON CORRIDOR

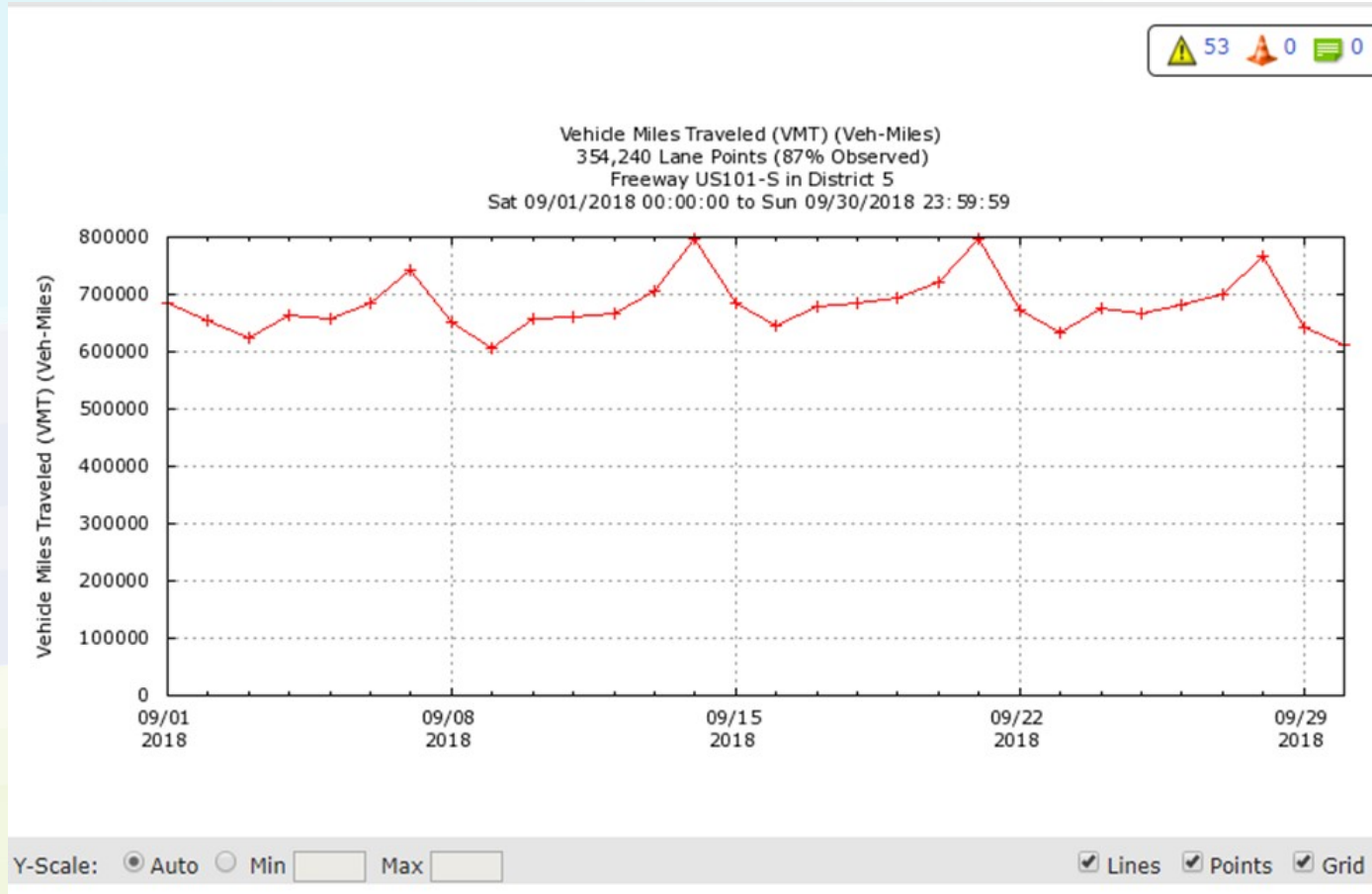
Objectives

- We want to assess traffic volumes for the US101 corridor.
 - » We want to see the VMT profile for the month of September
 - » We want to select traffic counts for a specific day in September
- We want to reduce the corridor section to only San Luis Obispo to Santa Maria
- We are interested in the SB direction

Process for VMT analysis

1. Select the corridor and geographic boundary of interest and select the 'Stations' report
2. Identify the section of interest mileposts
3. Navigate to Performance → Aggregates → Time Series
4. Modify dates for data and other parameters

VMT Output



Process for Day Counts

1. Select the corridor and geographic boundary of interest and select the 'Stations' report
2. Identify the section of interest mileposts
3. Navigate to Performance → Spatial Analysis → Multistation
4. Modify dates for data and other parameters to get the hourly counts for all stations in the area, for September 14.

Day Counts Output

⚠️ 3 🚦 0 📄 0

Hour	501014092	501014101	501015143	501016053	501016071	501016082	501016091	501016102	501016113	501016122	Data Quality	
	Mainline 89.834 (172.471)	Mainline 90.953 (173.590)	Mainline 14.705 (188.240)	Mainline 17.916 (191.451)	Mainline 19.797 (193.322)	Mainline 20.938 (194.463)	Mainline 21.995 (195.520)	Mainline 24.268 (197.793)	Mainline 24.806 (198.331)	Mainline 25.849 (199.366)	# Lane Points	% Observed
09/14/2018 00:00	294	344	241	287	271	301	260	259	250	221	276	75.0
09/14/2018 01:00	124	138	118	177	158	151	158	156	152	137	276	100.0
09/14/2018 02:00	116	133	111	138	134	138	144	142	140	115	276	100.0
09/14/2018 03:00	170	155	89	121	121	112	114	114	111	96	276	100.0
09/14/2018 04:00	292	261	119	127	137	129	136	138	135	116	276	100.0
09/14/2018 05:00	798	632	276	348	312	335	352	402	395	330	276	100.0
09/14/2018 06:00	1,724	1,539	752	934	881	898	918	980	965	759	276	100.0
09/14/2018 07:00	2,735	2,762	1,248	1,573	1,378	1,365	1,392	1,375	1,380	1,101	276	100.0
09/14/2018 08:00	2,384	2,287	1,282	1,779	1,473	1,487	1,522	1,590	1,574	1,184	276	100.0
09/14/2018 09:00	2,221	2,229	1,420	1,923	1,660	1,661	1,723	1,713	1,685	1,267	276	100.0
09/14/2018 10:00	2,284	2,324	1,555	2,198	1,956	1,921	1,976	1,963	1,962	1,491	276	100.0
09/14/2018 11:00	2,278	2,358	1,495	2,350	2,071	2,099	2,187	2,229	2,207	1,594	276	100.0
09/14/2018 12:00	2,316	2,469	1,465	2,653	2,414	2,424	2,495	2,503	2,478	1,918	276	100.0
09/14/2018 13:00	2,353	2,563	1,562	2,816	2,492	2,514	2,613	2,668	2,650	2,017	276	100.0
09/14/2018 14:00	2,656	2,998	1,862	3,313	3,016	3,035	3,129	3,099	3,051	2,418	276	100.0
09/14/2018 15:00	2,810	3,351	1,876	3,282	2,942	2,982	3,296	3,215	3,206	2,526	276	100.0
09/14/2018 16:00	2,839	3,571	2,340	3,281	2,877	2,775	3,352	3,346	3,349	2,681	276	100.0
09/14/2018 17:00	2,961	3,796	2,390	3,403	2,959	2,773	3,254	3,239	3,195	2,534	276	100.0
09/14/2018 18:00	2,955	3,476	1,559	2,899	2,569	2,436	2,512	2,577	2,546	1,979	276	100.0
09/14/2018 19:00	1,875	2,193	1,357	2,385	2,107	2,040	2,078	2,058	2,041	1,660	276	100.0
09/14/2018 20:00	1,496	1,701	1,147	1,844	1,706	1,641	1,653	1,705	1,683	1,392	276	100.0
09/14/2018 21:00	1,231	1,395	968	1,618	1,505	1,462	1,485	1,547	1,508	1,317	276	100.0
09/14/2018 22:00	998	1,127	915	1,424	1,304	1,298	1,348	1,339	1,316	1,150	276	100.0
09/14/2018 23:00	620	699	568	832	747	735	739	730	732	643	276	100.0
Total											6,624	99.0

NOTE: Showing first 10 of 19 stations in postmile range.

NOTE: To see data for all stations use 'Export to XLS'.

Let's try new queries!

- Can we get the average for all weekdays on a week?

EXERCISE 4 – EVALUATE TYPICAL SPEEDS ON CORRIDOR



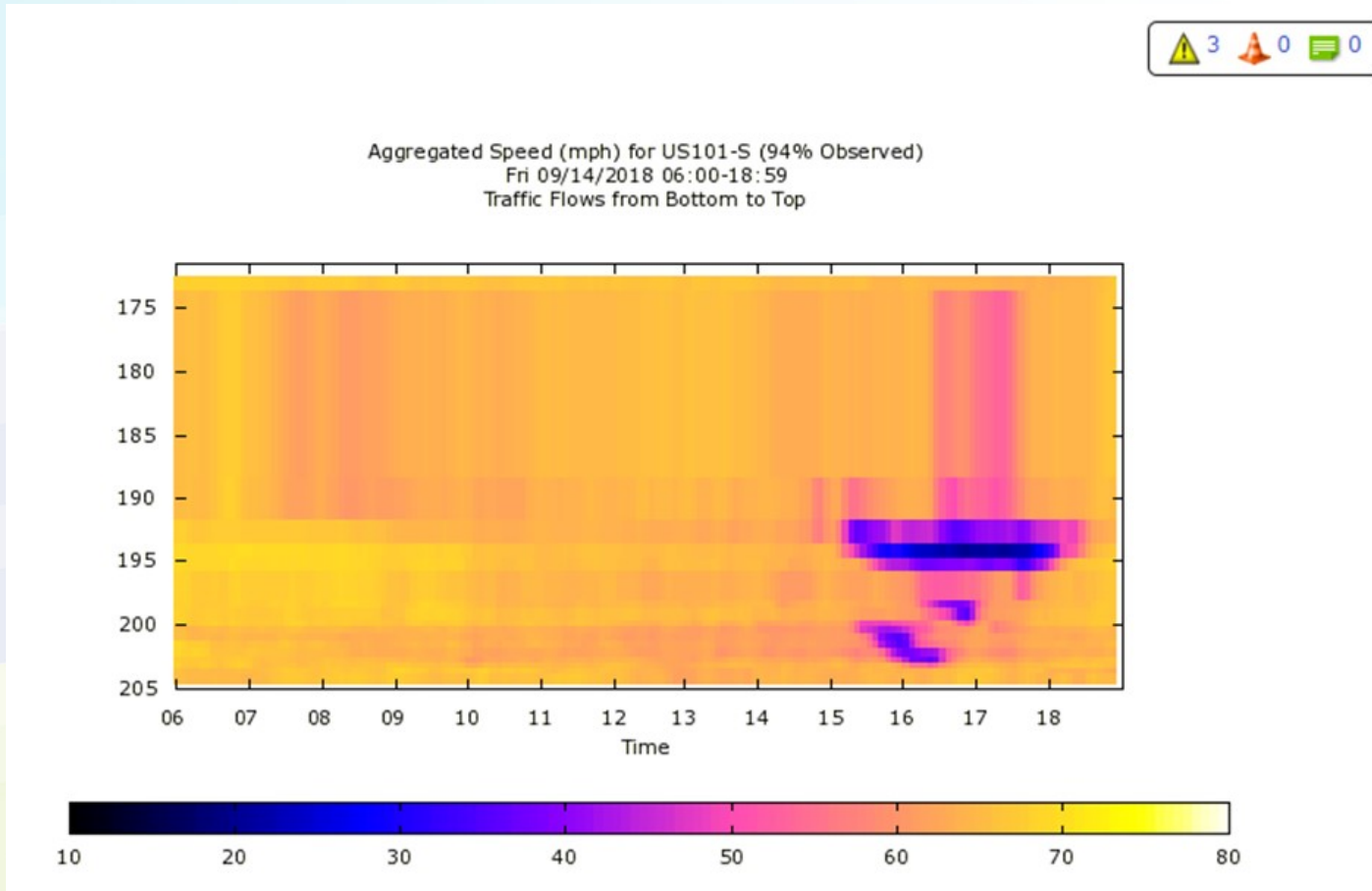
Objectives

- We want to assess speeds for the US101 corridor.
 - » We want to see speed contours and assess where are queues being formed on a particular day
 - » We want to see average speeds and assess where are queues being formed on an average weekday
- We want to reduce the corridor section to only San Luis Obispo to Santa Maria
- We are interested in the SB direction

Procedure for Speeds Contours on a Particular Day

1. Select the corridor and geographic boundary of interest and select the 'Stations' report
2. Identify the section of interest mileposts
3. Navigate to Performance → Spatial Analysis → Time Series Contours
4. Modify dates for data and other parameters to get the speeds for all stations in the area, for September 14.

Speed Contours for a Particular Day Outputs

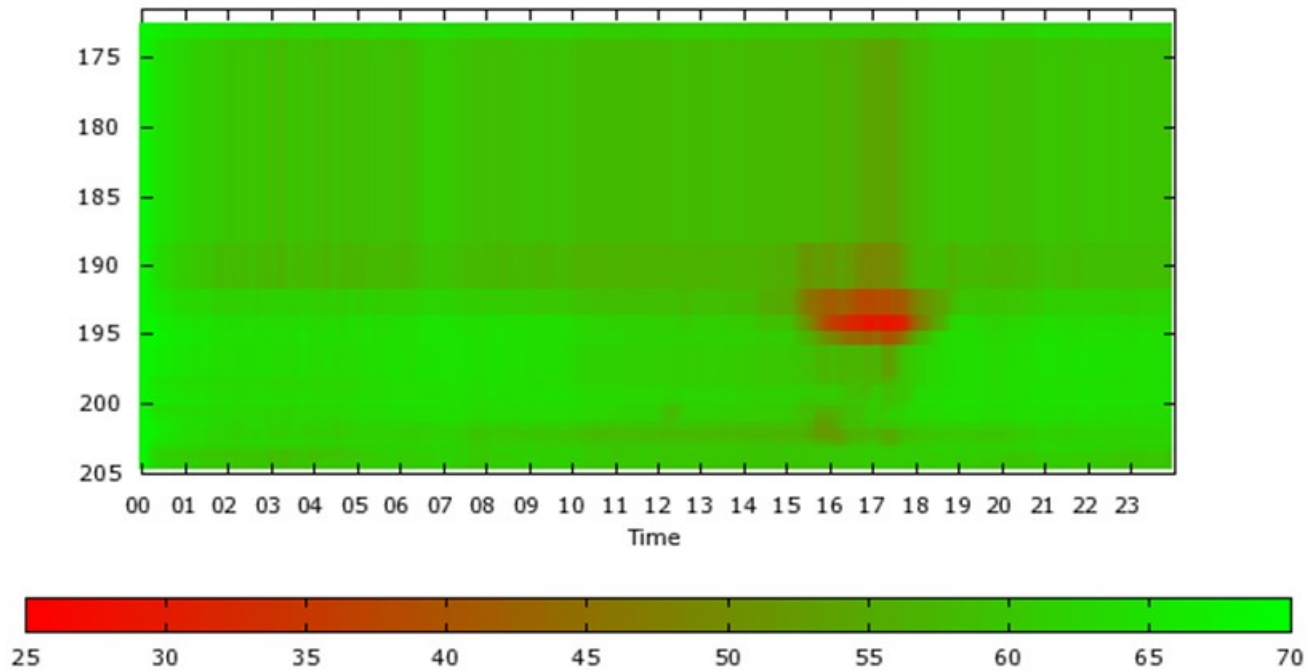


Procedure for Speeds Contours on a an Average Weekday

1. Select the corridor and geographic boundary of interest and select the 'Stations' report
2. Identify the section of interest mileposts
3. Navigate to Performance → Spatial Analysis → Long Contours
4. Modify dates for data and other parameters to get the speeds for all stations in the area, for September 14.

Speeds Contours on a an Average Weekday Outputs

Aggregated avg Weekday Speed (mph) for Sep 2018 (87% Observed)
District: 5, Segment Type: Freeway, Segment Name: US101-S
Traffic Flows from Bottom to Top



EXERCISE 5 – DOWNLOADING RAW DATA



PeMS Data Clearinghouse

The screenshot displays the PeMS 18.0 web interface for the State of California. The page is organized into several sections:

- Header:** Includes the PeMS 18.0 logo, a search bar, and navigation links for Home, Help, and Logout.
- Current Location:** A map of California with major cities like San Francisco, San Jose, Los Angeles, Tijuana, Ensenada, and Phoenix marked.
- Freeway Details:** A table listing key metrics:

Directional Distance	30,602.1 mi
Controllers	5,365
Stations	18,285
Detectors	45,088
Traffic Census Stations	16,527
- Status Check:** A 'Delay by Day of Week' line chart showing delay (veh-hrs) from Sunday to Saturday for 'Last Year', 'Last Week', and 'This Week'. The y-axis ranges from 0k to 2000k.
- Travel Time Reliability:** Two gauge charts showing reliability percentages for '5-10 AM' and '3-8 PM'.
- Detector Health:** A line chart showing '% Working' detectors over time, with a 'Statewide Goal: 90%' and 'Yesterday's Health: 67%'.
- Report Finder:** A form to search for reports by freeway, direction, and location.
- Announcements:** A list of recent updates, including 'Database Improvements' (September 12, 2018) and 'Database Maintenance' (September 04, 2018).
- Featured Sections:** A list of links to various reports and tools, with 'Data Clearinghouse' circled in red.

Downloading Data

- eMS Data Clearinghouse provides a single access point for downloading PeMS data sets.
- Data is available at 5-minutes, Hour, Day intervals.
- Station data for 5-minute intervals include:
 - » Timestamp
 - » Station ID
 - » District #
 - » Freeway #
 - » Direction of Travel
 - » Lane Type
 - » Station Length
 - » % of Observed Data
 - » Flow (by lane and total)
 - » Speed (by lane and total)
 - » Occupancy (by lane and total)
 - » Number of Data Samples (by lane)

Type
Station 5-Minute
District
District 7
Submit

D7 2018 Station 5-Minute

	J	F	M	A	M	J	J	A	S	O	N	D
18	█	█	█	█	█	█	█	█	█	█	█	█
17	█	█	█	█	█	█	█	█	█	█	█	█
16	█	█	█	█	█	█	█	█	█	█	█	█
15	█	█	█	█	█	█	█	█	█	█	█	█
14	█	█	█	█	█	█	█	█	█	█	█	█
13	█	█	█	█	█	█	█	█	█	█	█	█
12	█	█	█	█	█	█	█	█	█	█	█	█
11	█	█	█	█	█	█	█	█	█	█	█	█
10	█	█	█	█	█	█	█	█	█	█	█	█

Data Summary

This dataset contains the standard PeMS rollup of raw detector data. The algorithms used to process raw detector data are described in the System Help.

Months with data are indicated by a gray rectangle. Click a rectangle to view a listing of files available for download.

Field Specification

Name	Comment	Units
Timestamp	The date and time of the beginning of the summary interval. For example, a time of 08:00:00 indicates that the aggregate(s) contain measurements collected between 08:00:00 and 08:04:59. Note that second values are always 0 for five-minute aggregations. The format is MM/DD/YYYY HH24:MI:SS.	
Station	Unique station identifier. Use this value to cross-reference with <i>Metadata</i> files.	
District	District #	
Freeway	Freeway #	
Direction of Travel	N S E W	
Lane Type	A string indicating the type of lane. Possible values (and their meaning) are: <ul style="list-style-type: none"> • CD (Coll/Dist) • CH (Conventional Highway) • FF (Fwy-Fwy connector) • FR (Off Ramp) • HV (HOV) • ML (Mainline) • OR (On Ramp) 	
Station Length	Segment length covered by the station in miles/km.	
Samples	Total number of samples received for all lanes.	
%	Percentage of individual lane points at this	%

Available Files

File Name	Bytes
d07_text_station_5min_2018_01_01.txt.gz	29,746,960
d07_text_station_5min_2018_01_02.txt.gz	30,407,709
d07_text_station_5min_2018_01_03.txt.gz	30,085,444
d07_text_station_5min_2018_01_04.txt.gz	29,034,393
d07_text_station_5min_2018_01_05.txt.gz	30,545,474
d07_text_station_5min_2018_01_06.txt.gz	30,031,755
d07_text_station_5min_2018_01_07.txt.gz	28,880,544
d07_text_station_5min_2018_01_08.txt.gz	30,235,207
d07_text_station_5min_2018_01_09.txt.gz	29,396,188
d07_text_station_5min_2018_01_10.txt.gz	30,540,067
d07_text_station_5min_2018_01_11.txt.gz	30,779,357
d07_text_station_5min_2018_01_12.txt.gz	30,776,595
d07_text_station_5min_2018_01_13.txt.gz	30,194,909
d07_text_station_5min_2018_01_14.txt.gz	29,612,702
d07_text_station_5min_2018_01_15.txt.gz	30,337,947
d07_text_station_5min_2018_01_16.txt.gz	30,560,173
d07_text_station_5min_2018_01_17.txt.gz	30,519,709
d07_text_station_5min_2018_01_18.txt.gz	30,637,635
d07_text_station_5min_2018_01_19.txt.gz	30,579,904
d07_text_station_5min_2018_01_20.txt.gz	30,007,286
d07_text_station_5min_2018_01_21.txt.gz	28,598,466
d07_text_station_5min_2018_01_22.txt.gz	29,598,423
d07_text_station_5min_2018_01_23.txt.gz	29,486,038
d07_text_station_5min_2018_01_24.txt.gz	29,348,853

Example Sample

```
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
d07_text_station_5min_2018_09_15.txt
1 09/15/2018 00:00:00,715898,7,5,S,ML,.43,0,0,150,.0205,69.5,0,40,.0146,73.3,0,0,65,.0258,71,0,0,45,.0211,63.9,0,,,,,0,,,,,0,,,,,0,,,,,0
2 09/15/2018 00:00:00,715900,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
3 09/15/2018 00:00:00,715901,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
4 09/15/2018 00:00:00,715903,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
5 09/15/2018 00:00:00,715906,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
6 09/15/2018 00:00:00,715907,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
7 09/15/2018 00:00:00,715908,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
8 09/15/2018 00:00:00,715910,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
9 09/15/2018 00:00:00,715913,7,5,S,FR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
10 09/15/2018 00:00:00,715918,7,5,N,ML,1.075,0,0,154,.0275,69.5,0,41,.0196,73.3,0,0,66,.0346,71,0,0,47,.0283,63.9,0,,,,,0,,,,,0,,,,,0,,,,,0
11 09/15/2018 00:00:00,715920,7,5,S,ML,1.29,0,0,279,.0563,65,0,73,.0509,70.6,0,0,77,.0546,67.3,0,0,68,.059,62.1,0,0,61,.0606,58.9,0,,,,,0,,,,,0,,,,,0,,,,,0
12 09/15/2018 00:00:00,715923,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
13 09/15/2018 00:00:00,715924,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
14 09/15/2018 00:00:00,715925,7,5,S,OR,8,100,6,.0109,,8,6,.0109,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
15 09/15/2018 00:00:00,715926,7,5,N,OR,8,100,4,.0044,,8,4,.0044,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
16 09/15/2018 00:00:00,715927,7,5,N,OR,8,100,10,.0162,,8,10,.0162,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
17 09/15/2018 00:00:00,715928,7,5,S,OR,8,0,,8,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
18 09/15/2018 00:00:00,715929,7,5,S,ML,47,32,100,338,.0814,68.7,8,84,.0607,73.8,1,8,106,.1242,66.7,1,8,89,.0842,68.3,1,8,59,.0567,65.8,1,,,,,0,,,,,0,,,,,0,,,,,0
19 09/15/2018 00:00:00,715930,7,5,N,ML,505,32,100,327,.074,71.7,8,100,.0872,74.1,1,8,101,.0842,67.3,1,8,71,.0677,77.4,1,8,55,.0571,68.3,1,,,,,0,,,,,0,,,,,0,,,,,0
20 09/15/2018 00:00:00,715932,7,5,S,OR,8,100,15,.0194,,8,15,.0194,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
21 09/15/2018 00:00:00,715933,7,5,N,ML,425,32,100,286,.0544,70.9,8,106,.0722,72.6,1,8,86,.061,73.1,1,8,63,.0568,69,1,8,31,.0274,62.8,1,,,,,0,,,,,0,,,,,0,,,,,0
22 09/15/2018 00:00:00,715935,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
23 09/15/2018 00:00:00,715937,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
24 09/15/2018 00:00:00,715938,7,5,N,ML,545,32,100,349,.0775,65.7,8,109,.1076,67.3,1,8,118,.104,67.3,1,8,79,.0612,62.7,1,8,43,.0372,62.8,1,,,,,0,,,,,0,,,,,0,,,,,0
25 09/15/2018 00:00:00,715941,7,5,S,OR,8,100,5,.0209,,8,5,.0209,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
26 09/15/2018 00:00:00,715944,7,5,N,ML,323,32,100,272,.0662,70.2,8,81,.0688,76.4,1,8,89,.0846,70,1,8,71,.0797,66.6,1,8,31,.0319,62.8,1,,,,,0,,,,,0,,,,,0,,,,,0
27 09/15/2018 00:00:00,715947,7,5,S,ML,495,24,100,300,.0714,66.7,8,109,.0643,71.2,1,8,103,.0754,63.9,1,8,88,.0744,64.3,1,,,,,0,,,,,0,,,,,0,,,,,0
28 09/15/2018 00:00:00,715949,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
29 09/15/2018 00:00:00,715950,7,5,S,OR,8,0,,8,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
30 09/15/2018 00:00:00,715952,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
31 09/15/2018 00:00:00,715953,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
32 09/15/2018 00:00:00,715957,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
33 09/15/2018 00:00:00,715958,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
34 09/15/2018 00:00:00,715959,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
35 09/15/2018 00:00:00,715961,7,5,S,OR,10,100,16,.0253,,10,16,.0253,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
36 09/15/2018 00:00:00,715963,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
37 09/15/2018 00:00:00,715964,7,5,S,OR,10,100,12,.0211,,10,12,.0211,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
38 09/15/2018 00:00:00,715966,7,5,S,OR,10,100,2,.0023,,10,2,.0023,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
39 09/15/2018 00:00:00,715967,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
40 09/15/2018 00:00:00,715968,7,5,S,OR,10,100,18,.0176,,10,18,.0176,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
41 09/15/2018 00:00:00,715969,7,5,N,OR,10,100,2,.0021,,10,2,.0021,,1,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
42 09/15/2018 00:00:00,715970,7,5,N,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
43 09/15/2018 00:00:00,715971,7,5,S,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
44 09/15/2018 00:00:00,715972,7,210,E,OR,0,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0,,,,,0
```

Processing Raw Data

- Desired data is downloaded from the Clearinghouse, selecting District, interval (5-min, hour, or daily data), and the specific days.
- A Python script is used to read the Comma Separated Value data, and extract only the PeMS stations of interest.
- The script outputs an excel file with the desired data for further analysis.

```
In [1]: import pandas as pd
import os
import matplotlib.pyplot as plt
import numpy as np
%matplotlib inline
pd.options.display.max_columns = 50

In [2]: #We start by reading the PeMS stations we want to analyze
stations = pd.read_excel('pems_stations.xlsx', sheet_name='PeMS')

In [3]: #We now create a dataframe with all the speed data
head = ['Timestamp', 'Station', 'District', 'Route', 'Direction', 'Type', 'Length', 'Samples', '%Observed', 'TotalFlow', 'AvgOCc', 'AvgSpeed', 'Delay35', 'Delay40', 'Delay45', 'Delay50']
path = "Speeds/"

df = pd.DataFrame()

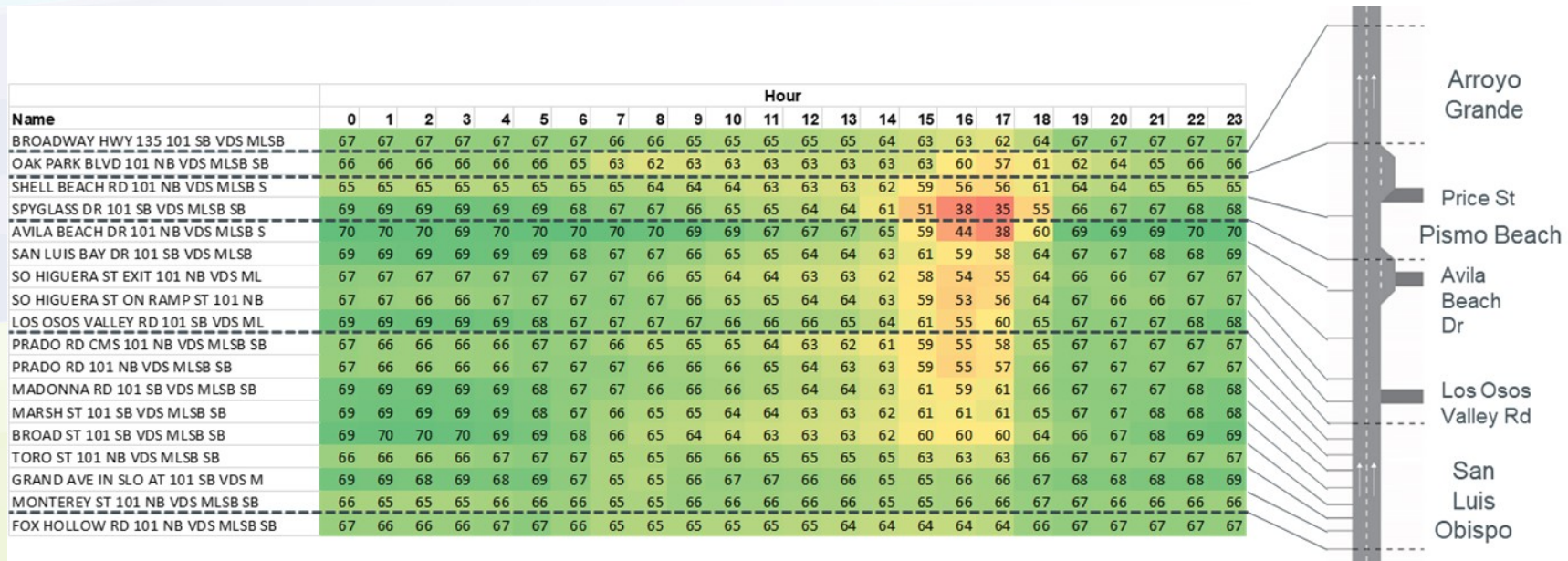
for file in os.listdir(path):
    filename = path+file
    temp = pd.read_csv(filename, header=None, names=head, parse_dates=['Timestamp'], infer_datetime_format=True)
    df = df.append(temp, ignore_index=True)

In [4]: df.head(10)
```

	Timestamp	Station	District	Route	Direction	Type	Length	Samples	%Observed	TotalFlow	AvgOCc	AvgSpeed	Delay35	Delay40	Delay45	Delay50
0	2017-01-01	500010011	5	1	N	ML	2.713	117	100	108.0	0.0115	64.2	0.0	0.0	0.0	0.0
1	2017-01-01	500010012	5	1	S	ML	2.714	117	100	41.0	0.0043	64.8	0.0	0.0	0.0	0.0
2	2017-01-01	500010021	5	1	N	ML	0.697	109	0	266.0	0.0226	66.9	0.0	0.0	0.0	0.0
3	2017-01-01	500010022	5	1	S	ML	0.698	109	0	171.0	0.0123	67.7	0.0	0.0	0.0	0.0
4	2017-01-01	500010031	5	1	N	ML	0.831	228	100	233.0	0.0134	68.1	0.0	0.0	0.0	0.0
5	2017-01-01	500010032	5	1	S	ML	0.831	114	100	202.0	0.0212	63.9	0.0	0.0	0.0	0.0
6	2017-01-01	500010041	5	1	N	ML	1.242	240	100	361.0	0.0182	68.0	0.0	0.0	0.0	0.0
7	2017-01-01	500010042	5	1	S	ML	1.241	240	100	220.0	0.0103	67.2	0.0	0.0	0.0	0.0
8	2017-01-01	500010072	5	1	S	ML	1.606	240	100	181.0	0.0066	66.6	0.0	0.0	0.0	0.0

Analyzing Raw Data

- Once data is selected for the region and time of interest, heat maps can be developed to visualize speeds, congestion, flows, and quality of data for a specific time period.



THANK YOU!

For further questions and comments you can email:

- Aldo Tudela – atudelarivadneyra@camsys.com
- Rajat Parashar – rparashar@camsys.com