

## Miguel\_Diaz\_ClassWork\_Sesson\_7

### SQL locally with PostgresSQL

#### Activity 1

Creating tables:

```
3 CREATE TABLE gps_data (  
4     datapointid VARCHAR(40) PRIMARY KEY,  
5     journeyid VARCHAR(40),  
6     latitude NUMERIC(10, 7),  
7     longitude NUMERIC(10, 7),  
8     month INTEGER,  
9     day INTEGER,  
10    hour INTEGER  
11 );  
12  
13 COPY gps_data (datapointid, journeyid, latitude, longitude, month, day, hour)  
14 FROM 'C:/adi10136/Documents/CE-5900-16/gps_data.csv'  
15 DELIMITER ','  
16 CSV HEADER;  
17
```

	datapointid [PK] character varying (40)	journeyid character varying (40)	latitude numeric (10,7)	longitude numeric (10,7)	month integer	day integer	hour integer
1	000074d0-2968-49ff-aa61-8f0ff648a3bd	80b68e1a324d2c47f413faef00cbf55033b6e73	41.5033700	-94.6438390	10	2	7
2	0000b449-4250-4b47-a4b9-f9c955db6cdd	9cc28b887c54bfd910e8f0c49cfe819122432d79	41.3631620	-94.6993880	10	1	6
3	000133f0-e2d7-446e-9036-8dccc1c11a74a	8cf32f521244f66a0795041ed458149ad245e996	41.4951030	-94.5879900	10	3	7
4	00017452-b680-4ef8-b84a-02d9b2d55e00	fdc51c3bf1171ad738d4d4180ed003b4372e02b9	41.3130080	-94.4531930	10	1	6
5	0002283c-b603-4cce-9521-c2c0d5f9cc14	21dfe4b9d65fe92c3833eca292b14618d33ac44d	41.4797720	-94.6829570	10	2	7
6	00026fca-6f10-48bb-bfa8-b65d85843c17	3dadf6e8e74bfbecbcb6551c27ac47c2e4e8d79f	41.4930800	-94.6671800	10	1	7
7	00029e52-a2a3-42c0-aecc-34f8bcd8439d	9ab0a042f7a696bf326cb68c4cce050544e63127	41.3019040	-94.5076160	10	2	2
8	000362e8-6a1e-46c3-9c18-a88d8110d1...	7447f426dba0f50f4491ac9b27d05b473c6bc64f	41.3019130	-94.3448860	10	3	6
9	00039125-10bf-4848-98bd-89835a030b90	d35bc46b290bce6d2f0dda65d3ac733d9b75b91c	41.4960780	-94.5104200	10	1	7
10	0003c83d-26ca-4f23-8de0-ccdbb5a441c1	aad413f86894c26134939e90bc528c3a4a49eb21	41.4931360	-94.2960720	10	2	7
11	00043376-1273-40eb-81d8-44e1766cec...	933189643d69d8a76afdd785cbe256aab4597a3a	41.2907940	-94.5626500	10	2	1
12	00044a58-335d-41d0-aeff-796288aa4e9a	1a1cc638f1d62d7428ef0d8117e04d6e0417cc3a	41.1862740	-94.4717440	10	1	7
13	0004b934-396d-4127-a531-daaa3f60cef4	b05102feef036405342947580880007d5d205c40	41.4968610	-94.5751210	10	1	7
14	0004dc3c-3a3d-48a0-b262-ec41f948c66b	bed5861d4e06cef32464704d743cd56886831703	41.4979560	-94.3176780	10	3	3
15	0005113f-da20-4887-b4b7-866a31d94297	47807fab870bd1775bf5218280a2d00922c16479	41.4929490	-94.2894780	10	1	7
16	00051863-9430-404c-931d-5e98e2d1c6...	79f272e334eccc7adce90ab541340b6db518156f	41.4930050	-94.3587250	10	3	5
17	00054749-a3ad-4d9f-bf3d-493aab8ead23	167458bea3f34cfe5969db1e4d9dd41d68dd1d45	41.4929800	-94.2630920	10	2	3
18	0006dd14-a24b-4c16-9150-c86460c962...	16ef58af500c1498c30ef52a0753cabb6c17d142	41.4968060	-94.5523840	10	2	2
19	0006f1de-e423-451a-aabd-4bc00981b219	52ec83558b66c61028a5ecb1f1628ce43d4b965f	41.4956930	-94.6429710	10	2	7
20	00072faf-d4b6-4284-a9fd-0710264d945a	7b6ec23ed88d5781fefb8632e523d87e8696d7ff	41.3742780	-94.5192590	10	1	7
21	0007c88f-4b42-41da-ba40-b6322aef406e	b7884bd1ad212cb20868c4512540d4d1476573...	41.4930620	-94.2500950	10	3	6
22	000812f5-d195-4e13-b9a5-12e66ddea10d	a2f9b7e63ba014946f2cd8e2a426102080c4a193	41.4969530	-94.5624010	10	2	7
23	00089760-60fe-40c2-91db-e329595f0516	78a049d6d9c86667d971d946373549aa3a1625...	41.4965330	-94.3191460	10	1	7
24	0008bad9-4429-4773-8b00-3effd6a5306a	f8e18d6a9d5e8a6e3ba3c5aaab7aea7588458658	41.2873810	-94.4796770	10	3	7

```

20 CREATE TABLE vehicle_data (
21     datapointid VARCHAR(40) PRIMARY KEY,
22     geohash VARCHAR(12),
23     speed NUMERIC(5, 2),
24     make VARCHAR(50),
25     model VARCHAR(50),
26     route_id VARCHAR(20),
27     segment_start_measure NUMERIC(10, 3)
28 );
29
30 COPY vehicle_data (datapointid, geohash, speed, make, model, route_id, segment_start_measure)
31 FROM 'C:/adi10136/Documents/CE-5900-16/vehicle_data.csv'
32 DELIMITER ','
33 CSV HEADER;

```

	datapointid [PK] character varying (40)	geohash character varying (12)	speed numeric (5,2)	make character varying (50)	model character varying (50)	route_id character varying (20)	segment_start_measure numeric (10,3)
1	000074d0-2968-49ff-aa61-8f0ff648a3bd	9zksn7	27.64	CHEVROLET	SILVERADO_HD	M002246110E	0.689
2	0000b449-4250-4b47-a4b9-f9c955db6cdd	9zke1	0.00	GMC	SIERRA_HD	C000141000N	9.110
3	000133f0-e2d7-446e-9036-8dcc1c11a74a	9zks9	122.11	GMC	SIERRA	S001910080E	78.961
4	00017452-b680-4ef8-b84a-02d9b2d55e00	9zkg48	61.05	GMC	ACADIA	S001930025N	48.283
5	0002283c-6b03-4cce-9521-c2c0d5f9cc14	9zkevr	97.91	CHEVROLET	TRAVERSE	C000141010N	1.254
6	00026fca-6f10-48bb-bfa8-b65d85843c17	9zksjc	124.41	GMC	SIERRA	S001910080E	74.798
7	00029e52-a2a3-42c0-aecc-34f8bcd8439d	9zkcfcq	86.39	GMC	TERRAIN	S001930092E	77.287
8	000362e8-6a1e-46c3-9c18-a88d8110d1...	9zksfn	94.46	CHEVROLET	SILVERADO_HD	S001930092E	85.787
9	00039125-10bf-4848-98bd-89835a030b90	9zku16	129.02	CHEVROLET	TRAX_TRACKER_TRAILBLAZER_B	S001910080E	82.973
10	0003c83d-26ca-4f23-8de0-cdbb5a441c1	9zkun1	118.65	CHEVROLET	SILVERADO_HD	S001910080W	211.620
11	00043376-1273-40eb-81d8-44e1766cec...	9zksfbh	9.21	CHEVROLET	EQUINOX	M264741070N	0.203
12	00044a58-335d-41d0-aeff-796288aa4e9a	9zksf61	0.00	CHEVROLET	SILVERADO	C000146050E	10.115
13	0004b934-396d-4127-a531-daea3f60cef4	9zkspf	116.35	CHEVROLET	EQUINOX	S001910080W	226.135
14	0004dc3c-3a3d-48a0-b262-ec41f948c66b	9zksjd	2.30	CHEVROLET	TRAVERSE	M745746070E	0.435
15	0005113f-da20-4887-b4b7-866a31d94297	9zksn3	119.80	CHEVROLET	IMPALA	S001910080E	94.501
16	00051863-9430-404c-931d-5e98e2d1c6...	9zksuhc	120.96	CHEVROLET	SILVERADO	S001910080E	90.930
17	00054749-a3ad-4d9f-bf3d-493aab8ead23	9zksunc	116.35	CHEVROLET	SILVERADO	S001910080E	95.896
18	0006dd14-a24b-4c16-9150-c86460c962...	9zksu06	118.65	GMC	TERRAIN	S001910080W	225.000

## Activity 2

### 1. Select datapoints from October 3 at 4 AM

```

36 SELECT *
37 FROM gps_data
38 WHERE month = 10 AND day = 3 AND hour = 4;
39

```

Data Output Messages Notifications

SQL

	datapointid [PK] character varying (40)	journeyid character varying (40)	latitude numeric (10,7)	longitude numeric (10,7)	month integer	day integer	hour integer
1	a5ed7547-1d79-4789-95b8-622f0ab95a...	118ddde885e9a0a6cce5618be206820180fa74...	41.4932230	-94.2579250	10	3	4
2	2204e9d5-1fde-4887-a842-62016eb95af4	118ddde885e9a0a6cce5618be206820180fa74...	41.4932200	-94.2591860	10	3	4
3	36c1a3c7-55d0-4ebf-a083-219f25dd85cf	118ddde885e9a0a6cce5618be206820180fa74...	41.4932190	-94.2604430	10	3	4
4	61f40a0b-4460-4142-9ada-2a1d293f337d	118ddde885e9a0a6cce5618be206820180fa74...	41.4932130	-94.2616950	10	3	4
5	c8270201-ee6b-4421-90fc-dae560bc2c55	118ddde885e9a0a6cce5618be206820180fa74...	41.4932070	-94.2629520	10	3	4
6	ed75bbad-feb6-4149-94cf-39831207e2c1	118ddde885e9a0a6cce5618be206820180fa74...	41.4932080	-94.2642110	10	3	4
7	992e15eb-d0ac-4650-9519-fba507823e...	118ddde885e9a0a6cce5618be206820180fa74...	41.4932050	-94.2654710	10	3	4
8	703f1f47-f773-422b-b2bf-b7a01b0648b8	118ddde885e9a0a6cce5618be206820180fa74...	41.4932050	-94.2667330	10	3	4
9	c372f3b6-6559-4467-bc8c-3e04d96928...	118ddde885e9a0a6cce5618be206820180fa74...	41.4931980	-94.2679960	10	3	4
10	112a08d3-ced1-4a7a-842b-ec5b72854c...	118ddde885e9a0a6cce5618be206820180fa74...	41.4931960	-94.2692570	10	3	4
11	c1757d70-b867-4843-9281-1fa7d1478b...	118ddde885e9a0a6cce5618be206820180fa74...	41.4931930	-94.2705170	10	3	4
12	3c5b48be-1f28-4470-991a-a4f9432b6a28	118ddde885e9a0a6cce5618be206820180fa74...	41.4931870	-94.2717780	10	3	4

2. Count how many rows in gps\_data table (Expected Answer: 244614)

```
41 SELECT COUNT(*)
42 FROM gps_data
43
44
```

Data Output		Messages	Notifications
	count bigint		
1	244614		

3. Find how many unique journey\_ids in the data filtered as per (1) (Ans: 7)

```
44 SELECT COUNT(DISTINCT journeyid)
45 FROM gps_data
46 WHERE month = 10 AND day = 3 AND hour = 4;
47
48
```

Data Output		Messages	Notifications
	count bigint		
1	7		

4. Filter data on October 1st and 3rd, and hour between 3 AM and 6 AM (Expected Answer: 101116 rows)

```
60 SELECT COUNT(*)
61 FROM gps_data
62 WHERE month = 10 AND day IN (1, 3);
63
64 SELECT hour, COUNT(*)
65 FROM gps_data
66 WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6
67 GROUP BY hour
68 ORDER BY hour;
69
70
71 SELECT *
72 FROM gps_data
73 WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6;
74
```

Data Output								Messages	Notifications
	datapointid [PK] character varying (40)	journeyid character varying (40)	latitude numeric (10,7)	longitude numeric (10,7)	month integer	day integer	hour integer		
1	757baad3-eec2-4cc1-9c9d-b1aef1d09a...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5037090	-94.3262400	10	1	5		
2	ba9d0a8f-8fe5-4bf8-8490-d70f3866b526	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5037060	-94.3259290	10	1	5		
3	4eb9fc5f-6541-4673-94af-d818e492c8e9	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5037000	-94.3255310	10	1	5		
4	82e68eb8-f6a2-4417-a317-6efc043fab88	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036960	-94.3250820	10	1	5		
5	d682dfbd-0304-4fe4-8859-3acbe4808a...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036950	-94.3246110	10	1	5		
6	7a546536-42fe-418e-aa5b-823373254a...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036870	-94.3241350	10	1	5		
7	d577364d-c36e-4151-8d6d-5fce79a3c8...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036820	-94.3236460	10	1	5		
8	ae9ed02e-cfa0-456c-addf-f579c998c056	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036830	-94.3231430	10	1	5		
9	29dc09a2-c357-42d6-8a3a-109858392c...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036860	-94.3226390	10	1	5		
10	36240474-41bb-4e0b-a0bf-3586cee433...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036890	-94.3221450	10	1	5		
11	ff7bb5f8-9414-4e15-bf75-ba4de558e179	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036850	-94.3216620	10	1	5		
12	ca6028dd-1601-4d0e-900a-f3dda4f1c2...	53ca9cce2fb74c2048a1f436c19e01b786086727	41.5036860	-94.3211810	10	1	5		
Total rows: 1000 of 74108    Query complete 00:00:00.344    Ln 74, Col 1									

```

60 SELECT COUNT(*)
61 FROM gps_data
62 WHERE month = 10 AND day IN (1, 3);
63
64 SELECT hour, COUNT(*)
65 FROM gps_data
66 WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6
67 GROUP BY hour
68 ORDER BY hour;
69
70
71 -- SELECT *
72 -- FROM gps_data
73 -- WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6;
74

```

Data Output Messages Notifications

	hour integer	count bigint
1	3	5715
2	4	10578
3	5	13249
4	6	44566

5. Select all journey\_ids containing “abc” from the whole gps\_data table (Expected Answer: 8)

```

76 SELECT DISTINCT journeyid
77 FROM gps_data
78 WHERE journeyid LIKE '%abc%';
79
80

```

Data Output Messages Notifications

	journeyid character varying (40)
1	209abc451e58c092f42a5d95945593e8a5419...
2	2a68a21131f7796346598f5aedbab62f573eba...
3	4aabc63593bd9e85c4e6718905c245204a9c2...
4	6cabcb55bfda791d53b0b27c2dcad6e859b9ca...
5	7abc0b68ce38a1942369073329d44bf43079a...
6	896e29a3ea3e270540d1c6cab9bfff5f099e62d
7	c97784415e2893ad6eeac5c9ebabc4b26ca23...
8	cdb490fc11874d45d929d0b01f2abc6f19c031af

6. Count how many datapoint\_ids start with “9” (Expected Answer: 15256)

```

81 SELECT COUNT(*)
82 FROM gps_data
83 WHERE datapointid LIKE '9%';
84

```

Data Output Messages Notifications

	count bigint
1	15256

## 7. Select all rows where journey\_id ends with “cc” (Expected Answer: 66 rows)

```

85 SELECT *
86 FROM gps_data
87 WHERE journeyid LIKE '%cc';
88

```

Data Output Messages Notifications

	datapointid [PK] character varying (40)	journeyid character varying (40)	latitude numeric (10,7)	longitude numeric (10,7)	month integer	day integer	hour integer
1	50630098-4c43-4b83-ad1c-b783af1ab7...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4963860	-94.3181120	10	3	5
2	e7bce1a8-578b-4324-af62-f7e1b57e14f3	ca5acfb22db167611af25761f2f677dee1f622cc	41.4963860	-94.3181120	10	3	5
3	fbf58fac-45ce-422e-a81c-ab0cc4328664	ca5acfb22db167611af25761f2f677dee1f622cc	41.4963860	-94.3181120	10	3	5
4	cf5e6f6f-35e0-4169-9dfc-42f6c9c292cb	ca5acfb22db167611af25761f2f677dee1f622cc	41.4963860	-94.3181120	10	3	5
5	18e21731-6207-418f-b3a1-6e65beaae6f4	ca5acfb22db167611af25761f2f677dee1f622cc	41.4963860	-94.3181110	10	3	5
6	e773d360-4079-40c6-ac5d-a7448ea85b...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4963890	-94.3180760	10	3	5
7	7e9c1b7a-7123-4440-a875-35aeb9259...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4964300	-94.3180070	10	3	5
8	6005c8cc-9060-4d84-bd09-3655a761af...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4965300	-94.3179740	10	3	5
9	f90990c2-2f83-4b9e-9c5c-0d49524829eb	ca5acfb22db167611af25761f2f677dee1f622cc	41.4966640	-94.3179750	10	3	5
10	fa920e21-1e92-4901-85a9-d03a711079...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4967610	-94.3181050	10	3	5
11	2c7acd01-a7ba-4eee-b432-c49e69f17c...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4967870	-94.3182750	10	3	5
12	b5a24832-f18c-4266-b9c0-2d1230b73f...	ca5acfb22db167611af25761f2f677dee1f622cc	41.4967710	-94.3184240	10	3	5

Total rows: 66 of 66 Query complete 00:00:00.144 Ln 84, Col 1

## 8. Select all columns and create a new column that combines datapoint\_id and day with a “-” symbol

```

89 SELECT *, CONCAT(datapointid, '-', day) AS dpid_day
90 FROM gps_data;
91

```

Data Output Messages Notifications

	datapointid [PK] character varying (40)	journeyid character varying (40)	latitude numeric (10,7)	longitude numeric (10,7)	month integer	day integer	hour integer	dpid_day text
1	757baad3-eec2-4cc1-9c9d-b1aef1d09acc	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5037090	-94.3262400	10	1	5	757baad3-eec2-4cc1-9c9d-b1aef1d09acc-1
2	bafd0a8f-8fe5-4fbf-8490-d70f3866b526	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5037060	-94.3259290	10	1	5	bafd0a8f-8fe5-4fbf-8490-d70f3866b526-1
3	4eb9fc5f-6541-4673-94af-d818e492c8e9	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5037000	-94.3255310	10	1	5	4eb9fc5f-6541-4673-94af-d818e492c8e9-1
4	82e68eb8-f6a2-4417-a317-6efc043fab88	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036960	-94.3250820	10	1	5	82e68eb8-f6a2-4417-a317-6efc043fab88-1
5	d682dfbd-0304-4fe4-8859-3acbe4808a2a	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036950	-94.3246110	10	1	5	d682dfbd-0304-4fe4-8859-3acbe4808a2a-1
6	7a546536-42fe-418e-aa5b-823373254a...	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036870	-94.3241350	10	1	5	7a546536-42fe-418e-aa5b-823373254a87-1
7	d577364d-c36e-4151-8d6d-5fce79a3c8e3	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036820	-94.3236460	10	1	5	d577364d-c36e-4151-8d6d-5fce79a3c8e3-1
8	ae9ed02e-cfa0-456c-addf-f579c998c056	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036830	-94.3231430	10	1	5	ae9ed02e-cfa0-456c-addf-f579c998c056-1
9	29dc09a2-c357-42d6-8a3a-109858392c...	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036860	-94.3226390	10	1	5	29dc09a2-c357-42d6-8a3a-109858392c54-1
10	36240474-41bb-4e0b-a0bf-3586cee433...	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036890	-94.3221450	10	1	5	36240474-41bb-4e0b-a0bf-3586cee433c9-1
11	ff7bb5f8-9414-4e15-bf75-ba4de558e179	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036850	-94.3216620	10	1	5	ff7bb5f8-9414-4e15-bf75-ba4de558e179-1
12	c4b028dd-1601-4d0e-900a-f3dda4f1c2aa	53ca9cce2fb74c2048a1f436c19e01b7860867...	41.5036860	-94.3211810	10	1	5	c4b028dd-1601-4d0e-900a-f3dda4f1c2aa-1

Total rows: 3000 of 244614 Query complete 00:00:00.545 Ln 82, Col 17

### Activity – 3: Aggregations

1. Find maximum speed from table vehicle\_data (Ans: 162.43)

```
100 SELECT MAX(speed) AS max_speed
101 FROM vehicle_data;
102
103
104
```

Data Output Messages Notifications



	max_speed numeric
1	162.43

2. Find average, minimum, maximum and standard deviations of speeds for “CHEVROLET” make

```
103 SELECT
104     AVG(speed) AS avg_speed,
105     MIN(speed) AS min_speed,
106     MAX(speed) AS max_speed,
107     STDDEV(speed) AS stddev_speed
108 FROM vehicle_data
109 WHERE make = 'CHEVROLET';
110
```

Data Output Messages Notifications



	avg_speed numeric	min_speed numeric	max_speed numeric	stddev_speed numeric
1	79.3176807801335612	0.00	157.82	47.0256958229966512

3. Create seg\_id by combining “route\_id” and “segment\_start\_measure”

```
111 SELECT
112     route_id || '-' || segment_start_measure AS seg_id,
113     *
114 FROM vehicle_data;
115
116
```

Data Output Messages Notifications



	seg_id text	datapointid [PK] character varying (40)	geohash character varying (12)	speed numeric (5,2)	make character varying (50)	model character varying (50)	route_id character varying (20)	segment_start_measure numeric (10,3)
1	C000141010N-2.218	a36a9bb9-1fc8-4d07-bbef-2923988cb99a	9zksjb	97.91	CHEVROLET	TRAVERSE	C000141010N	2.218
2	C000141010N-2.293	8bfc8779-1029-4af4-9414-5beddc4c5807	9zksjb	97.91	CHEVROLET	TRAVERSE	C000141010N	2.293
3	C000141010N-2.344	45412676-a4a1-4458-aad6-5eb860dda6...	9zksjb	95.61	CHEVROLET	TRAVERSE	C000141010N	2.344
4	C000141010N-2.344	4ef14e98-a6c8-4ae0-9215-57a9d3cdda37	9zksjb	91.00	CHEVROLET	TRAVERSE	C000141010N	2.344
5	C000141010N-2.429	e34b206b-0785-458c-81f1-c0abe28fb6e3	9zksjb	88.70	CHEVROLET	TRAVERSE	C000141010N	2.429
6	C000141010N-2.457	e56076e7-2306-4936-9c6c-96a4921c1f...	9zksjb	88.70	CHEVROLET	TRAVERSE	C000141010N	2.457
7	C000141060N-12.408	d42ad9b5-6850-4475-9682-8f12ccf2a737	9zksjb	89.85	CHEVROLET	TRAVERSE	C000141060N	12.408
8	C000141060N-12.481	0f968847-693c-4f75-8459-3da14a6f8cea	9zksjc	97.91	CHEVROLET	TRAVERSE	C000141060N	12.481
9	C000141060N-12.522	9e21912c-8309-4b40-84b5-bc3dd0141a...	9zksjc	99.07	CHEVROLET	TRAVERSE	C000141060N	12.522
10	C000141060N-12.522	fcee60e3-4456-46f4-befc-f197fcb4f92b	9zksjc	99.07	CHEVROLET	TRAVERSE	C000141060N	12.522
11	C000141060N-12.601	0effc69e-1c72-4cd4-bf92-767cd026f84b	9zksjc	99.07	CHEVROLET	TRAVERSE	C000141060N	12.601
12	C000141060N-12.655	4cfa3849-2747-4e30-a62a-309707be3a...	9zksjc	99.07	CHEVROLET	TRAVERSE	C000141060N	12.655

4. Add this seg\_id to existing vehicle\_data table and verify it
- Alter\_table and update

```
111 SELECT
112     route_id || '-' || segment_start_measure AS seg_id,
113     *
114 FROM vehicle_data;
115
116 ALTER TABLE vehicle_data
117 ADD COLUMN seg_id VARCHAR(60);
118
119 UPDATE vehicle_data
120 SET seg_id = route_id || '-' || segment_start_measure;
121
122 SELECT seg_id, *
123 FROM vehicle_data
124 LIMIT 10;
125
```

Data Output Messages Notifications									
	jointid	geohash	speed	make	model	route_id	segment_start_measure	seg_id	
	character varying (40)	character varying (12)	numeric (5,2)	character varying (50)	character varying (50)	character varying (20)	numeric (10,3)	character varying (60)	
1	74d0-2968-49ff-aa61-8f0ff648a3bd	9zksn7	27.64	CHEVROLET	SILVERADO_HD	M002246110E	0.689	M002246110E-0.689	
2	b449-4250-4b47-a4b9-f9c955db6cdd	9zkem1	0.00	GMC	SIERRA_HD	C000141000N	9.110	C000141000N-9.110	
3	33f0-e2d7-446e-9036-8dccc1c11a74a	9zksps9	122.11	GMC	SIERRA	S001910080E	78.961	S001910080E-78.961	
4	7452-b680-4ef8-b84a-02d9b2d55e00	9zkg48	61.05	GMC	ACADIA	S001930025N	48.283	S001930025N-48.283	
5	283c-6b03-4cce-9521-c2c0d5f9cc14	9zkevr	97.91	CHEVROLET	TRAVERSE	C000141010N	1.254	C000141010N-1.254	
6	6fca-6f10-48bb-bfa8-b65d85843c17	9zksjc	124.41	GMC	SIERRA	S001910080E	74.798	S001910080E-74.798	
7	9e52-a2a3-42c0-aecc-34f8bcd8439d	9zkcfc	86.39	GMC	TERRAIN	S001930092E	77.287	S001930092E-77.287	
8	62e8-6a1e-46c3-9c18-a88d8110d1...	9zktvn	94.46	CHEVROLET	SILVERADO_HD	S001930092E	85.787	S001930092E-85.787	
9	9125-10bf-4848-98bd-89835a030b90	9zku16	129.02	CHEVROLET	TRAX_TRACKER_TRAILBLAZER_B	S001910080E	82.973	S001910080E-82.973	
10	c83d-26ca-4f23-8de0-cdcb5a441c1	9zkun1	118.65	CHEVROLET	SILVERADO_HD	S001910080W	211.620	S001910080W-211.620	
11	3376-1273-40eb-81d8-44e1766cec...	9zktbh	9.21	CHEVROLET	EQUINOX	M264741070N	0.203	M264741070N-0.203	
12	4a58-335d-41d0-aeff-796288aa4e9a	9zkt61	0.00	CHEVROLET	SILVERADO	C000146050F	10.115	C000146050F-10.115	
Total rows: 244614 of 244614 Query complete 00:00:01.948 Ln 1, Col 1									

5. Find 85th percentiles speeds for each segment\_id

```
125
126 SELECT
127     seg_id,
128     percentile_cont(0.85) WITHIN GROUP (ORDER BY speed) AS percentile_85_speed
129 FROM vehicle_data
130 GROUP BY seg_id;
131
```

Data Output Messages Notifications		
	seg_id	percentile_85_speed
	character varying (60)	double precision
1	C000141000N-2.997	13.469
2	C000141000N-3.011	35.477
3	C000141000N-3.061	42.62
4	C000141000N-3.086	48.035000000000004
5	C000141000N-3.161	48.38
6	C000141000N-3.198	44.92
7	C000141000N-3.241	42.735
8	C000141000N-3.320	39.972
9	C000141000N-3.375	31.617499999999996
10	C000141000N-3.453	26.49
11	C000141000N-3.497	22.923999999999996
12	C000141000N-9.110	0
Total rows: 1000 of 8745 Query complete 00:00:00.994 Ln 125, Col 1		



## 6. Find the highest speeds for each model? (Compare 95 th percentile speeds for each model)

```

132 SELECT
133     model,
134     MAX(speed) AS max_speed
135 FROM vehicle_data
136 GROUP BY model;
137
138 SELECT
139     model,
140     percentile_cont(0.95) WITHIN GROUP (ORDER BY speed) AS percentile_95_speed
141 FROM vehicle_data
142 GROUP BY model;
143

```

Data Output Messages Notifications

	model character varying (50)	percentile_95_speed double precision
1	ACADIA	126.71
2	ATS	118.65
3	AVEO_SONIC	130.17
4	BLAZER	126.71
5	CAMARO	127.87
6	CANYON	80.64
7	COLORADO	129.02
8	CORVETTE_STINGRAY	124.41
9	CRUZE	129.02
10	CTS	127.87
11	ENCLAVE	123.26
12	ENCORE	136.4474999999998

Total rows: 34 of 34 Query complete 00:00:00.408 Ln 130, Col 20

## ACTIVITY - 4 : Joins

### 1. Join the “gps\_data” and “vehicle\_data” tables using “JOIN” on datapoint\_id

```

148 SELECT *
149 FROM gps_data
150 JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid;
151
152

```

Data Output Messages Notifications

	datapointid character varying (40)	journeyid character varying (40)	latitude numeric (10,7)	longitude numeric (10,7)	month integer	day integer	hour integer	datapointid character varying (40)	geohash character varying (10)
1	c38c9038-bd14-47b7-8b4b-a2ff0dd355...	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	41.4959820	-94.6611990	10	1	2	c38c9038-bd14-47b7-8b4b-a2ff0dd355...	9zksjf
2	65a98c70-1e88-4a4d-a3cb-3d04a224e...	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	41.4967920	-94.6586050	10	1	2	65a98c70-1e88-4a4d-a3cb-3d04a224e...	9zksjf
3	5f027be4-a200-44a7-a011-4900fb66e1...	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	41.4978010	-94.6501130	10	1	2	5f027be4-a200-44a7-a011-4900fb66e1...	9zksn4
4	a61bf214-cd6e-47e9-878c-d73c4e5560...	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	41.4977910	-94.6451930	10	1	2	a61bf214-cd6e-47e9-878c-d73c4e5560...	9zksn6
5	e6d6dde4-2c95-4b0d-a5c3-1c4916635...	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	41.4977840	-94.6436760	10	1	2	e6d6dde4-2c95-4b0d-a5c3-1c4916635...	9zksn6
6	45fc686d-8dbb-4f9e-9317-f6c5f68bc449	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	41.5021250	-94.6434030	10	1	2	45fc686d-8dbb-4f9e-9317-f6c5f68bc449	9zksn7
7	fa5858fc-35e5-4020-9f2f-6bc7598527ed	3e743ba1b7241b76315fe8aa489311d2304d7d...	41.4931430	-94.2679520	10	1	2	fa5858fc-35e5-4020-9f2f-6bc7598527ed	9zkunc
8	67eed3b7-b732-46c4-909c-0e05793ca2...	3e743ba1b7241b76315fe8aa489311d2304d7d...	41.4931410	-94.2691930	10	1	2	67eed3b7-b732-46c4-909c-0e05793ca2...	9zkunc
9	e4ccd255-1c77-4473-93d7-d6545900a...	3e743ba1b7241b76315fe8aa489311d2304d7d...	41.4931280	-94.2754030	10	1	2	e4ccd255-1c77-4473-93d7-d6545900a...	9zkun9
10	52fa7fa4-00ea-47e8-bc41-3ddbe669a9...	3e743ba1b7241b76315fe8aa489311d2304d7d...	41.4931060	-94.2853390	10	1	2	52fa7fa4-00ea-47e8-bc41-3ddbe669a9...	9zkun3
11	d149700c-d422-4535-890c-1491ba589...	3e743ba1b7241b76315fe8aa489311d2304d7d...	41.4930870	-94.2915280	10	1	2	d149700c-d422-4535-890c-1491ba589...	9zkun3
12	h1a9a7f5-a8fa-4796-h090-h01afa01188	3e743ba1b7241b76315fe8aa489311d2304d7d...	41.4930830	-94.2938380	10	1	2	h1a9a7f5-a8fa-4796-h090-h01afa01188	9zkun1

Total rows: 1000 of 244614 Query complete 00:00:01.135 Ln 151, Col 1



		datapointid character varying (40)	geohash character varying (12)	speed numeric (5,2)	make character varying (50)	model character varying (50)	route_id character varying (20)	segment_start_measure numeric (10,3)	seg_id character varying (60)
1	2	130c70303001447107040421100033...	9zksj1	97.91	CHEVROLET	TRAVERSE	C000141060N	1.2.073	C000141060N-1.2.073
2	2	65a98c70-1e88-4a4d-a3cb-3d04a224e...	9zksj1	93.31	CHEVROLET	TRAVERSE	C000146580E	1.185	C000146580E-1.185
3	2	5f027be4-a200-44a7-a011-4900fb66e1...	9zksn4	46.08	CHEVROLET	TRAVERSE	M002246040E	0.384	M002246040E-0.384
4	2	a61bf214-cd6e-47e9-978c-d73c4e5560...	9zksn6	35.71	CHEVROLET	TRAVERSE	M002246040E	0.614	M002246040E-0.614
5	2	e6d6dde4-2c95-4b0d-a5c3-1c4916635...	9zksn6	13.82	CHEVROLET	TRAVERSE	M002246040E	0.724	M002246040E-0.724
6	2	45fc686d-8d8b-4f9e-9317-f6c5f68bc449	9zksn7	47.23	CHEVROLET	TRAVERSE	M002241060N	0.830	M002241060N-0.830
7	2	fa5858fc-35e5-4020-9f2f-6bc7598527ed	9zkunc	123.26	GMC	SIERRA	S001910080W	210.156	S001910080W-210.156
8	2	67eed3b7-b732-46c4-909c-0e05793ca2...	9zkunc	123.26	GMC	SIERRA	S001910080W	210.261	S001910080W-210.261
9	2	e4ccd255-1c77-4473-93d7-d6545900a...	9zkun9	124.41	GMC	SIERRA	S001910080W	210.567	S001910080W-210.567
10	2	52fa7fa4-00ea-47e8-bc41-3ddb669a9...	9zkun3	123.26	GMC	SIERRA	S001910080W	211.115	S001910080W-211.115
11	2	d149700c-d422-4535-890c-1a91ba589...	9zkun3	123.26	GMC	SIERRA	S001910080W	211.414	S001910080W-211.414
12	2	b1e9e2f5-e8da-4796-b909-b91a6a0188...	9zkun1	123.26	GMC	SIERRA	S001910080W	212.064	S001910080W-212.064

2. Select datapoint\_id, journey\_id, speed, seg\_id from the above join

```

147
148 SELECT *
149 FROM gps_data
150 JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid;
151
152 SELECT gps_data.datapointid, gps_data.journeyid, vehicle_data.speed, vehicle_data.seg_id
153 FROM gps_data
154 JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid;
155

```

Data Output Messages Notifications

	datapointid character varying (40)	journeyid character varying (40)	speed numeric (5,2)	seg_id character varying (60)
1	9e21912c-8309-4b40-84b5-bc3dd0141a...	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	99.07	C000141060N-12.522
2	fcee60e3-4456-46f4-befc-f197fcb4f92b	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	99.07	C000141060N-12.522
3	c38c9038-bd14-47b7-8b4b-a2ff0dd35543	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	97.91	C000141060N-12.895
4	65a98c70-1e88-4a4d-a3cb-3d04a224e262	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	93.31	C000146580E-1.185
5	29f64967-6cf8-4f04-97a9-c65a80c84d92	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	86.39	M002246040E-0.000
6	5f027be4-a200-44a7-a011-4900fb66e148	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	46.08	M002246040E-0.384
7	58bf14a0-051c-4d35-9a21-6eab37fbd1b6	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	46.08	M002246040E-0.462
8	92ea6cb5-8d9c-422b-9103-c7b63f3016d3	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	46.08	M002246040E-0.462
9	3e96f2f0-5365-4201-8513-321cece9cbaf	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	33.40	M002246040E-0.542
10	02c1e5f5-6403-4595-ae9-63cf834d7f21	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	29.95	M002246040E-0.570
11	88615f74-6c34-4044-998a-0246b541e677	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	36.86	M002246040E-0.614
12	a61bf214-cd6e-47e9-978c-d73c4e55604b	e74cc1aca5d0d95b4ae48f5086d4f6ec1c6159...	35.71	M002246040E-0.614

Total rows: 1000 of 244614 Query complete 00:00:01.610 Ln 142, Col 19

3. Count the unique journey\_ids that came to stop at least once (Speed = 0)

```

155
156 SELECT COUNT(DISTINCT gps_data.journeyid) AS stopped_journey_count
157 FROM gps_data
158 JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid
159 WHERE vehicle_data.speed = 0;
160

```

Data Output Messages Notifications

	stopped_journey_count bigint
--	---------------------------------

1	646
---	-----

ANEXUS (IMPLEMENTED CODE):

-- ACTIVITY 1 - CREATE TABLES

-- DROP TABLE IF EXISTS gps\_data;

-- CREATE TABLE gps\_data (

-- datapointid VARCHAR(40) PRIMARY KEY,

-- journeyid VARCHAR(40),

-- latitude NUMERIC(10, 7),

-- longitude NUMERIC(10, 7),

-- month INTEGER,

-- day INTEGER,

-- hour INTEGER

-- );

-- COPY gps\_data (datapointid, journeyid, latitude, longitude, month, day, hour)

-- FROM 'C:/adi10136/Documents/CE-5900-16/gps\_data.csv'

-- DELIMITER ','

-- CSV HEADER;

-- DROP TABLE IF EXISTS vehicle\_data;

-- CREATE TABLE vehicle\_data (

-- datapointid VARCHAR(40) PRIMARY KEY,

-- geohash VARCHAR(12),

-- speed NUMERIC(5, 2),

-- make VARCHAR(50),

-- model VARCHAR(50),

-- route\_id VARCHAR(20),

-- segment\_start\_measure NUMERIC(10, 3)

-- );

-- COPY vehicle\_data (datapointid, geohash, speed, make, model, route\_id, segment\_start\_measure)

-- FROM 'C:/adi10136/Documents/CE-5900-16/vehicle\_data.csv'

-- DELIMITER ','

-- CSV HEADER;

-----

-- ACTIVITY 2

```

-- SELECT *

-- FROM gps_data

-- WHERE month = 10 AND day = 3 AND hour = 4;


-- SELECT COUNT(*)

-- FROM gps_data


-- SELECT COUNT(DISTINCT journeyid)

-- FROM gps_data

-- WHERE month = 10 AND day = 3 AND hour = 4;


-- -- Query to select the rows

-- SELECT *

-- FROM gps_data

-- WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6;


-- -- Query to count the rows that match the condition

-- SELECT COUNT(*)

-- FROM gps_data

-- WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6;


-- SELECT COUNT(*)

-- FROM gps_data

-- WHERE month = 10 AND day IN (1, 3);


-- SELECT hour, COUNT(*)

-- FROM gps_data

-- WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6

-- GROUP BY hour

-- ORDER BY hour;


-- -- SELECT *

-- -- FROM gps_data

-- -- WHERE month = 10 AND day IN (1, 3) AND hour BETWEEN 3 AND 6;


-- SELECT DISTINCT journeyid

-- FROM gps_data

-- WHERE journeyid LIKE '%abc%';

```

```

-- SELECT COUNT(*)

-- FROM gps_data

-- WHERE datapointid LIKE '9%';


-- SELECT *

-- FROM gps_data

-- WHERE journeyid LIKE '%cc';


-- SELECT *, CONCAT(datapointid, '-', day) AS dpid_day

-- FROM gps_data;

-----

-- ACTIVITY - 3 - Aggregations


-- SELECT MAX(speed) AS max_speed

-- FROM vehicle_data;


-- SELECT

--   AVG(speed) AS avg_speed,

--   MIN(speed) AS min_speed,

--   MAX(speed) AS max_speed,

--   STDDEV(speed) AS stddev_speed

-- FROM vehicle_data

-- WHERE make = 'CHEVROLET';


-- SELECT

--   route_id || '-' || segment_start_measure AS seg_id,

--   *

-- FROM vehicle_data;


-- ALTER TABLE vehicle_data

-- ADD COLUMN seg_id VARCHAR(60);


-- UPDATE vehicle_data

-- SET seg_id = route_id || '-' || segment_start_measure;


-- SELECT seg_id, *

-- FROM vehicle_data

-- LIMIT 10;


-- SELECT

--   seg_id,

--   percentile_cont(0.85) WITHIN GROUP (ORDER BY speed) AS percentile_85_speed

-- FROM vehicle_data

```

```
-- GROUP BY seg_id;
```

```
-- SELECT
```

```
--   model,
```

```
--   MAX(speed) AS max_speed
```

```
-- FROM vehicle_data
```

```
-- GROUP BY model;
```

```
-- SELECT
```

```
--   model,
```

```
--   percentile_cont(0.95) WITHIN GROUP (ORDER BY speed) AS percentile_95_speed
```

```
-- FROM vehicle_data
```

```
-- GROUP BY model;
```

-----

```
-- ACTIVITY - 4 : Joins
```

```
-- SELECT *
```

```
-- FROM gps_data
```

```
-- JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid;
```

```
-- SELECT gps_data.datapointid, gps_data.journeyid, vehicle_data.speed, vehicle_data.seg_id
```

```
-- FROM gps_data
```

```
-- JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid;
```

```
-- SELECT COUNT(DISTINCT gps_data.journeyid) AS stopped_journey_count
```

```
-- FROM gps_data
```

```
-- JOIN vehicle_data ON gps_data.datapointid = vehicle_data.datapointid
```

```
-- WHERE vehicle_data.speed = 0;
```