Determine home and work locations from user GPS, mobile and wifi data

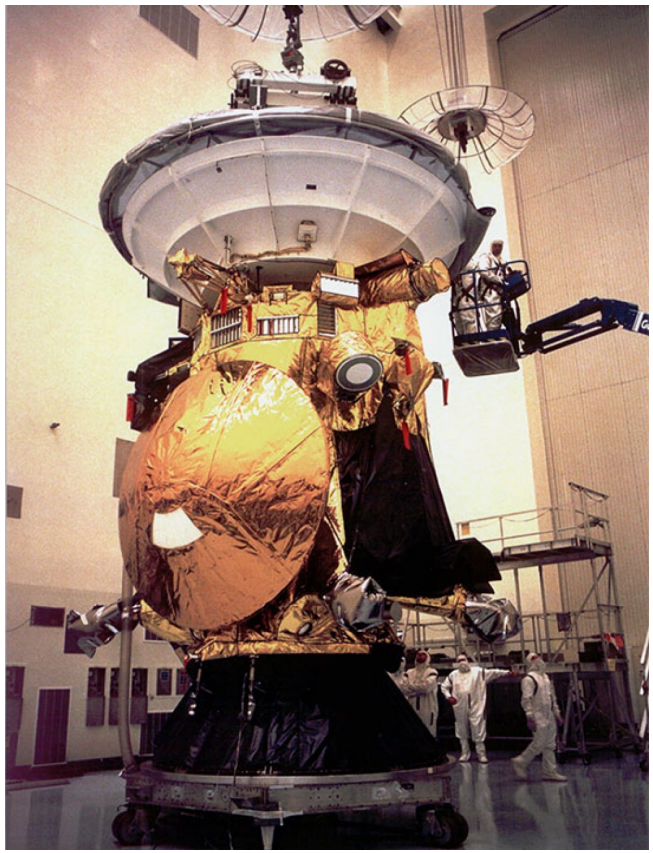
Shengyi Ye

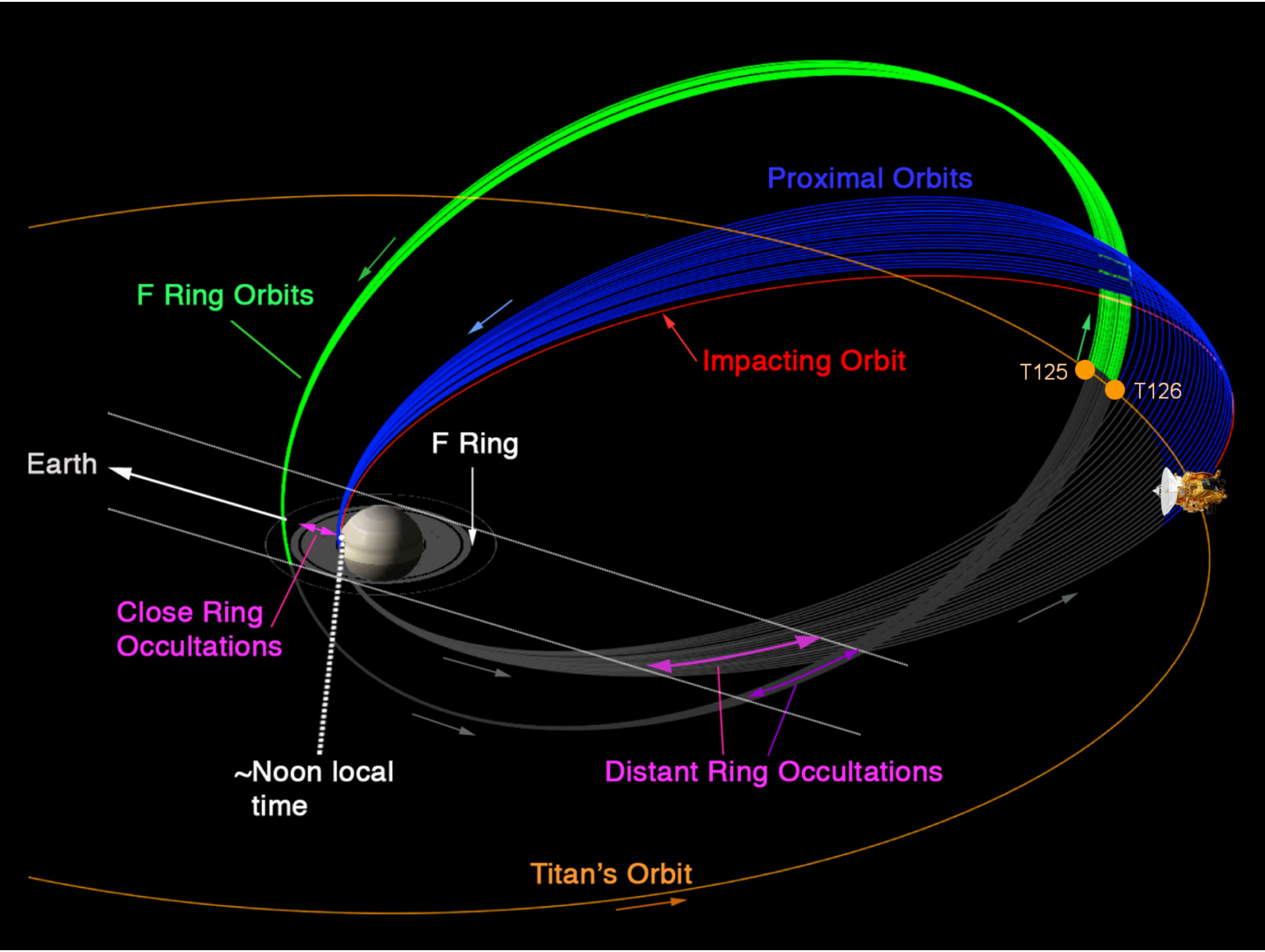
My background

space physics

Cassini spacecraf

NAIF/spice kerna

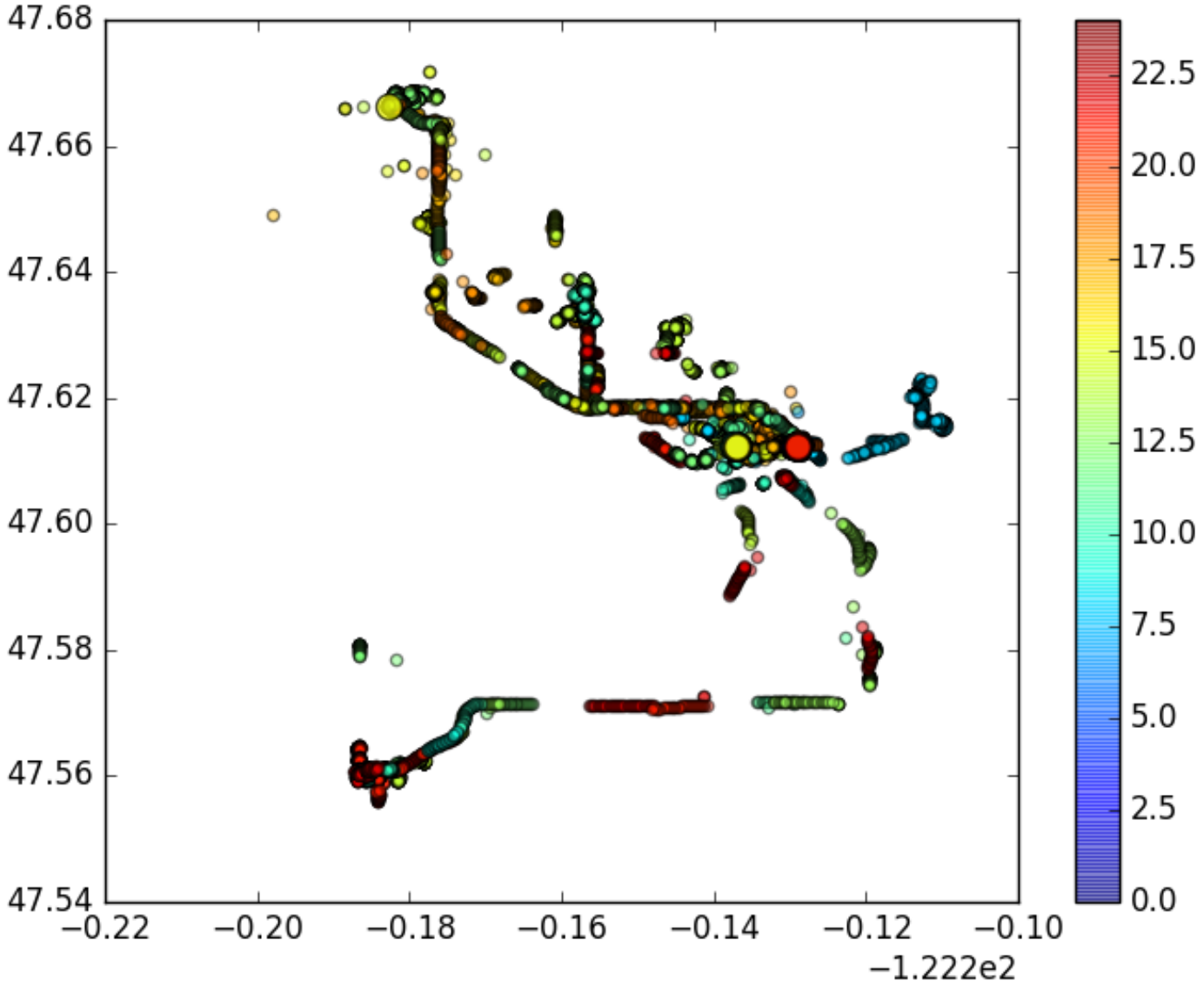



Proximal Orbits
F Ring Orbits
Impacting Orbit
T125
T126
F Ring
Earth
Close Ring
Occultations
~Noon local
time
Distant Ring Occultations
Titan's Orbit

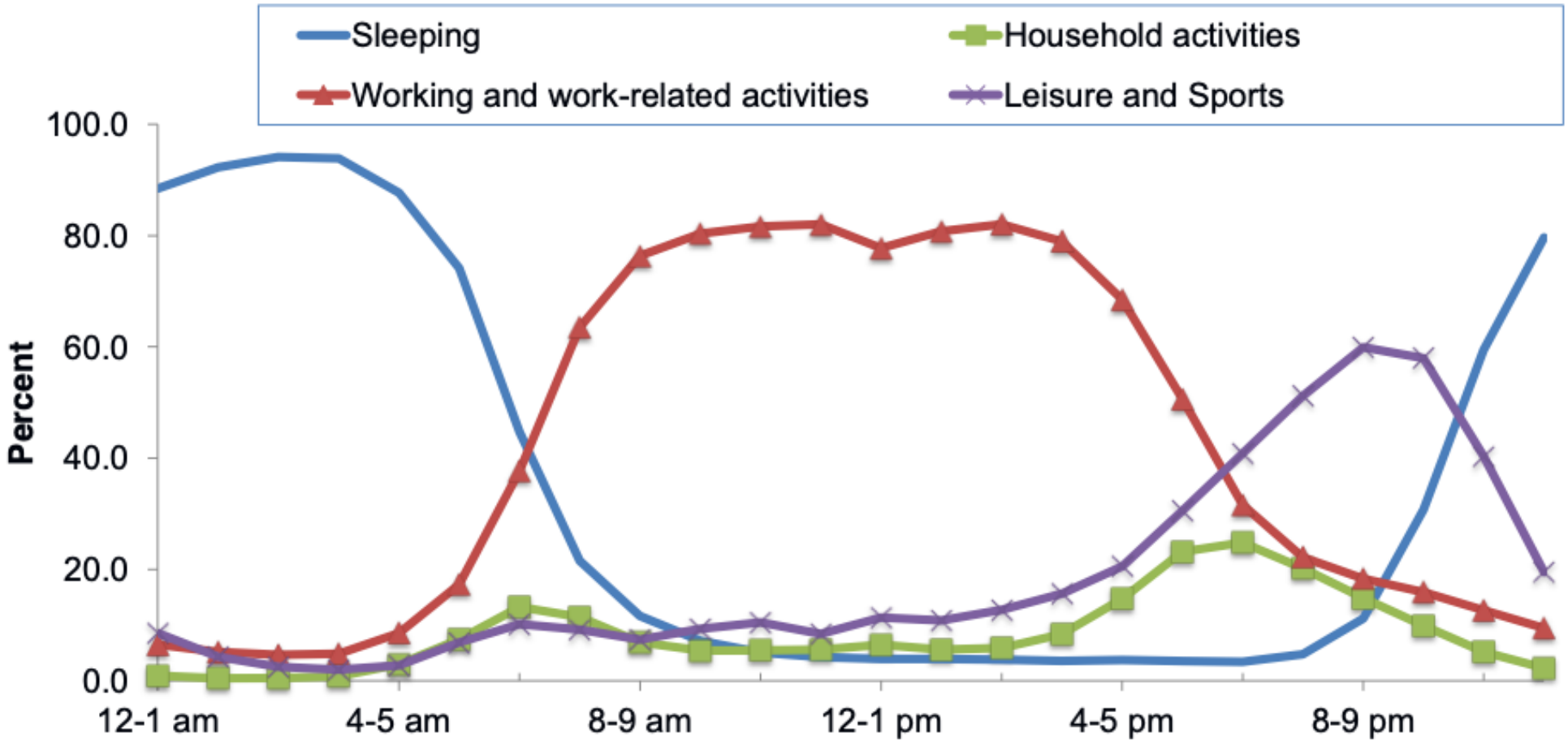
Steps

- Visualize locations
- Compute local times
- Grid data by time
- Reduce data
- Clustering
- Determine home and work locations

User 2 (5935)


47.68
47.66
47.64
47.62
47.60
47.58
47.56
47.54
−0.22
−0.20
−0.18
−0.16
−0.14
−0.12
−0.10
−1.222e2
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

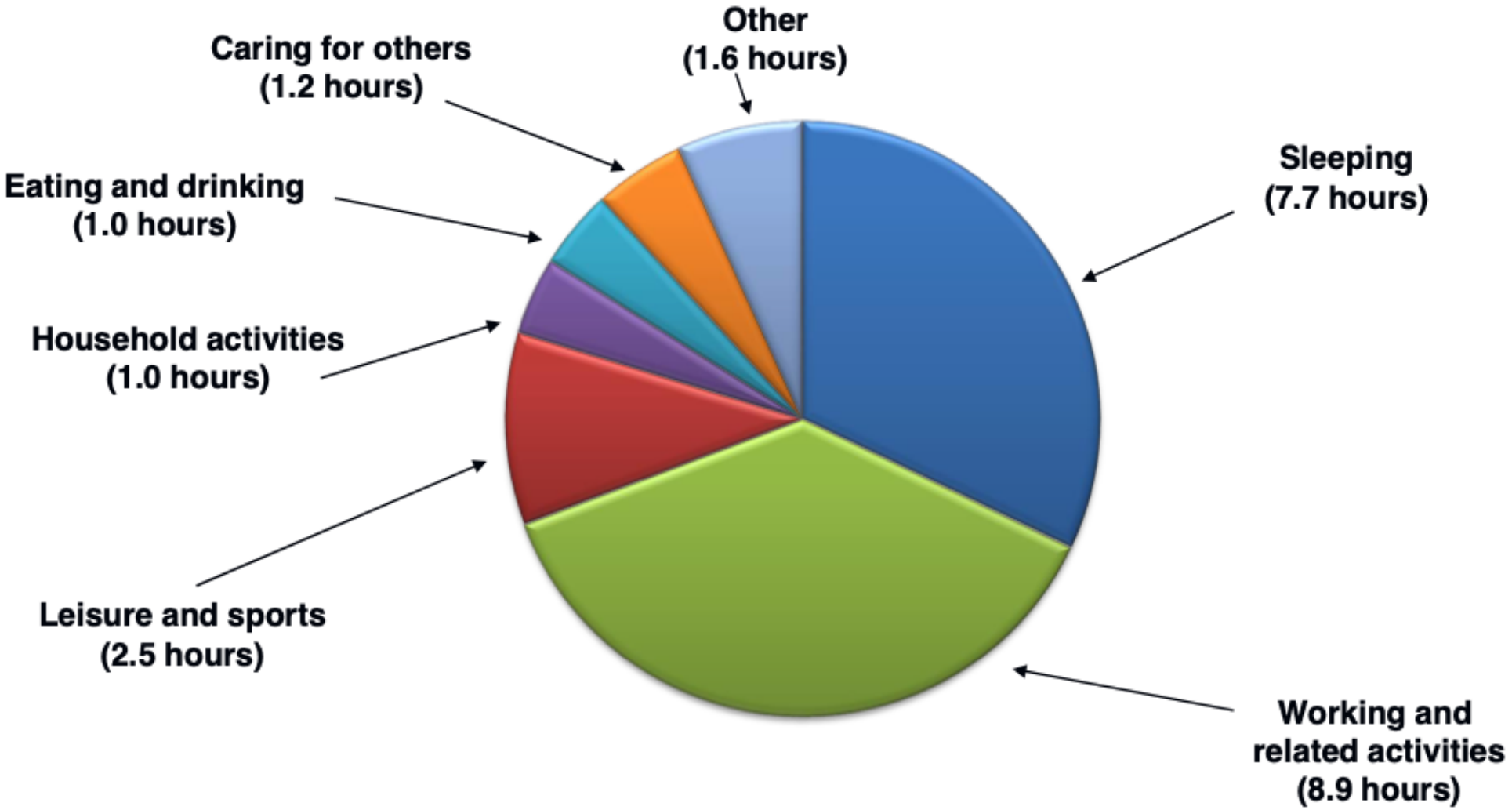
Percent of employed persons who did selected activities on workdays by hour of the day



NOTE: Data include individuals, age 15 and over, who were employed full-time on days they worked. Data are an average for 2010-14.

SOURCE: Bureau of Labor Statistics, American Time Use Survey

Time use on an average work day for employed persons ages 25 to 54 with children



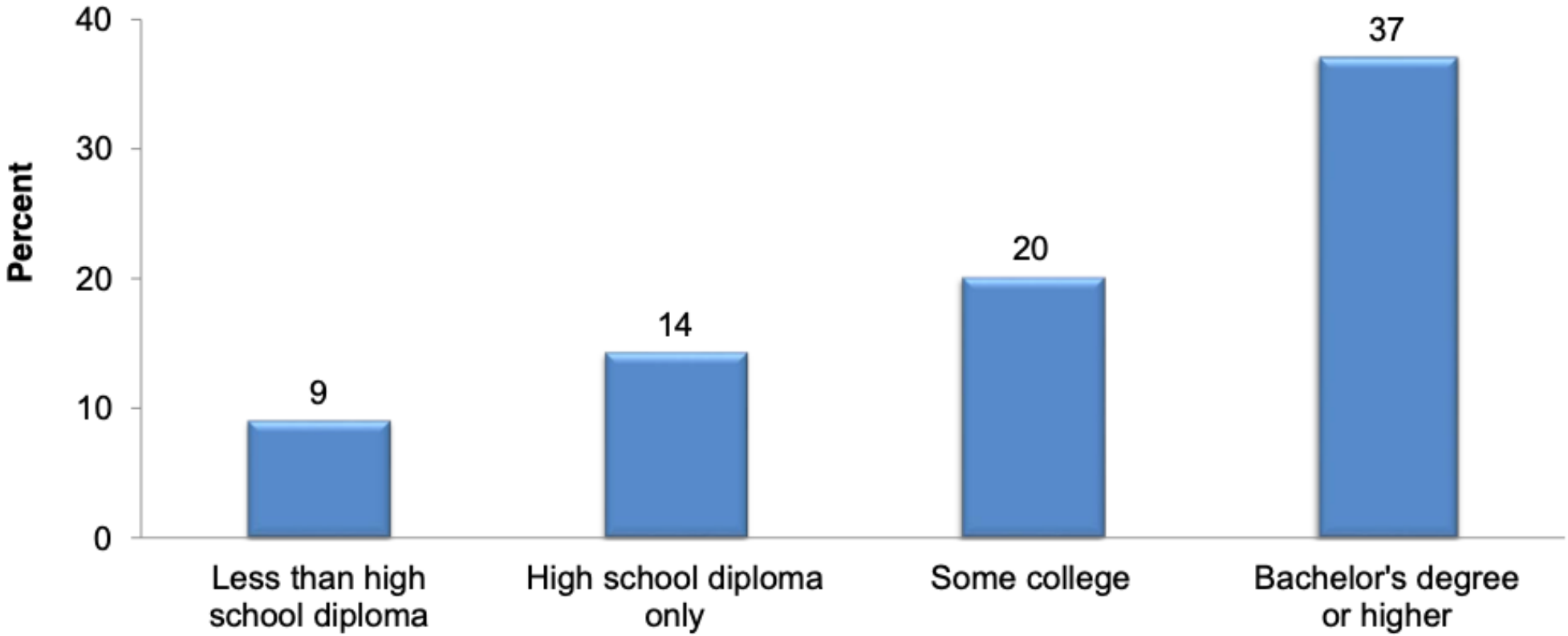
NOTE: Data include employed persons on days they worked, ages 25 to 54, who lived in households with children under 18. Data include non-holiday weekdays and are annual averages for 2014. Data include related travel for each activity.

SOURCE: Bureau of Labor Statistics, American Time Use Survey

Assumptions

- One home: most time spent
- One job: 2nd most time spent daytime
- Normal work schedule: 8-18
- Occasional travel
- Places: VAR(location) < 1e-7 within one hour

Percent of employed persons who worked at home on an average workday, by education level



NOTE: Data include all employed persons age 25 and over on days they worked. Working at home includes any time persons did work at home and is not restricted to persons whose usual workplace is their home. Data include all days of the week and are averages for 2010-14.

SOURCE: Bureau of Labor Statistics, American Time Use Survey

Simplify

- Use longitude to estimate local time

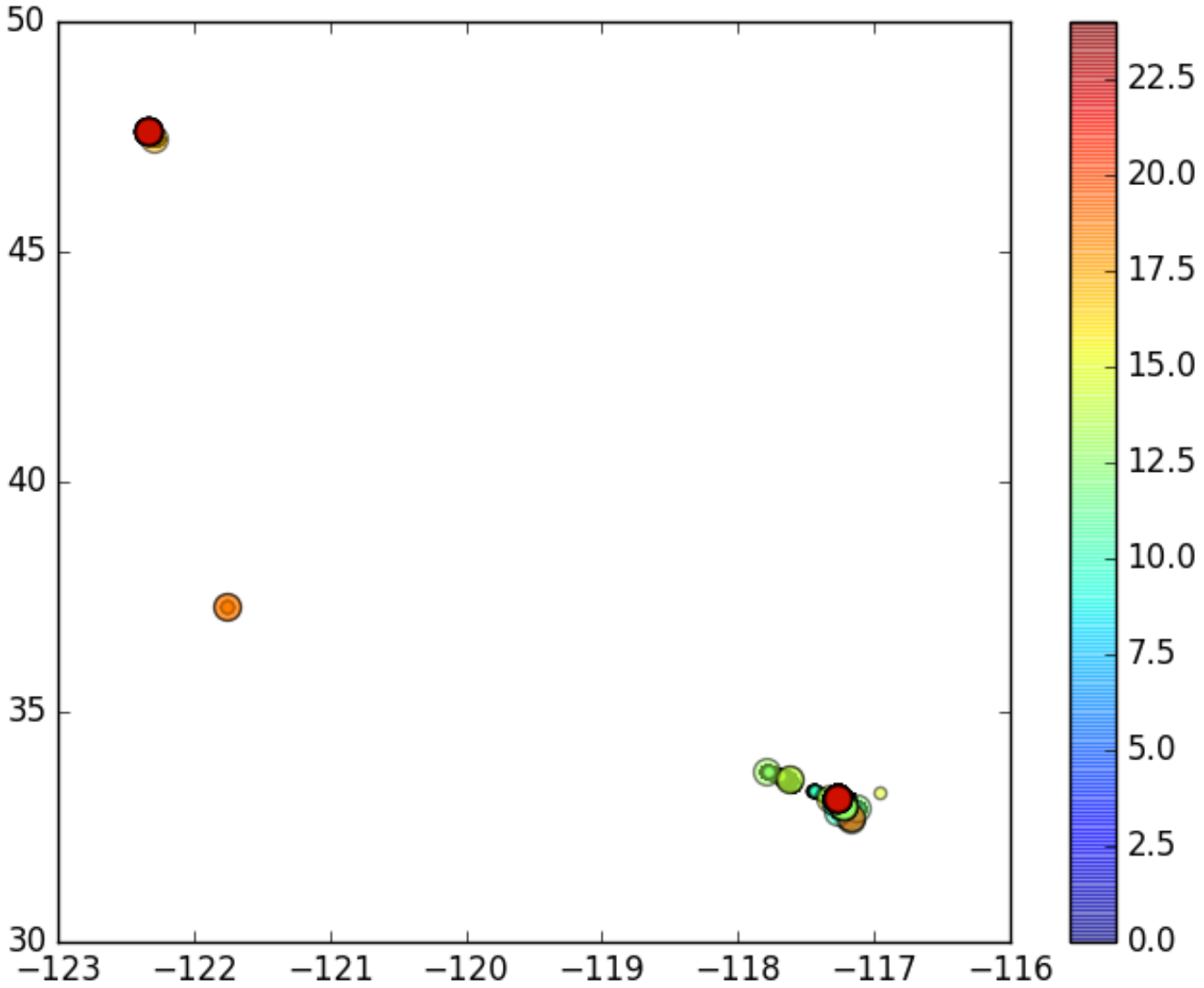
  Δh= longitude/15°
- Longitude and latitude as 2D coordinates
- Calculate distance and average locations
- Ignore circular effect, assume person don’t travel over the globe (tricky if longitude near +/-180°, circular average needed)

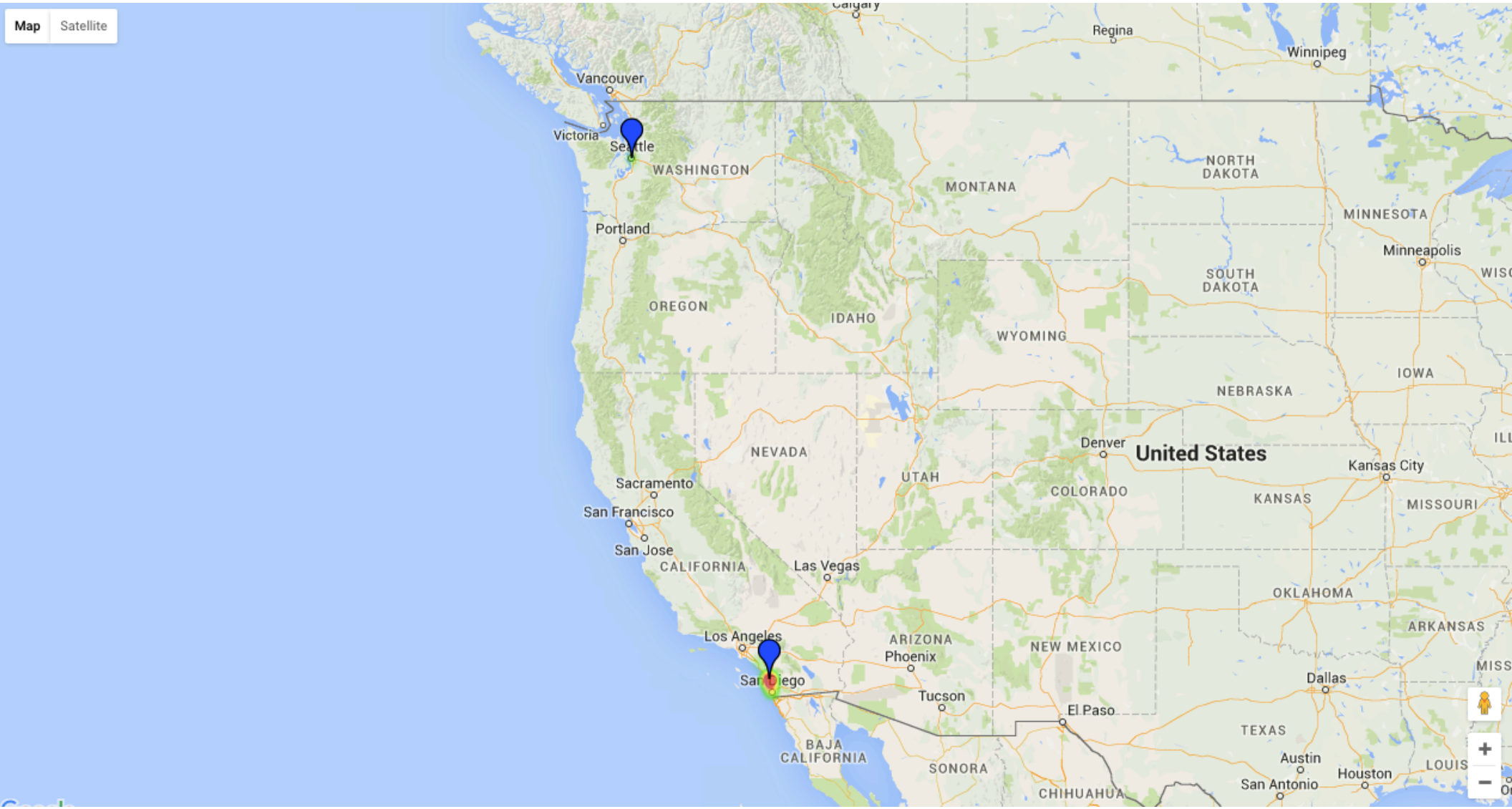
User 1

2016-01-07
Thursday

(-74.01247690000001, 40.716731199999998)	0	2.01948391737e-28
(-74.01247690000001, 40.716731199999998)	1	2.01948391737e-28
(-74.01247690000001, 40.716731199999991)	2	2.52435489671e-28
(-74.01247690000001, 40.716731199999998)	8	2.01948391737e-28
(-74.005131800000001, 40.742140750000004)	10	2.1497125003e-09
(-74.005125242857147, 40.742139814285721)	11	2.10584081662e-09
(-74.005065500000015, 40.742114177777779)	12	1.09514061733e-08
(-74.005696174999997, 40.742193512500002)	13	5.2262061047e-07
(-74.005032350000008, 40.742081633333335)	14	1.99368280564e-08
(-74.005122619999995, 40.742108600000002)	16	5.00929600097e-10
(-74.005125242857147, 40.742128800000003)	17	1.77431102066e-09
(-74.005164585714283, 40.742134414285715)	18	8.45065306211e-10
(-74.01247690000001, 40.716731199999991)	20	2.52435489671e-28
(-74.01247690000001, 40.716731199999998)	21	2.01948391737e-28

User 0


50
45
40
35
30
−123
−122
−121
−120
−119
−118
−117
−116
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0


Map
Satellite
Regina
Winnipeg
Vancouver
Victoria
Seattle
WASHINGTON
MONTANA
NORTH DAKOTA
MINNESOTA
Minneapolis
Portland
SOUTH DAKOTA
OREGON
IDAHO
WYOMING
IOWA
NEBRASKA
NEVADA
Denver
United States
Kansas City
Sacramento
UTAH
COLORADO
KANSAS
MISSOURI
San Francisco
San Jose
CALIFORNIA
Las Vegas
OKLAHOMA
ARKANSAS
Los Angeles
ARIZONA
Phoenix
NEW MEXICO
San Diego
Dallas
Tucson
El Paso
TEXAS
BAJA CALIFORNIA
SONORA
Austin
Houston
San Antonio
CHIHUAHUA

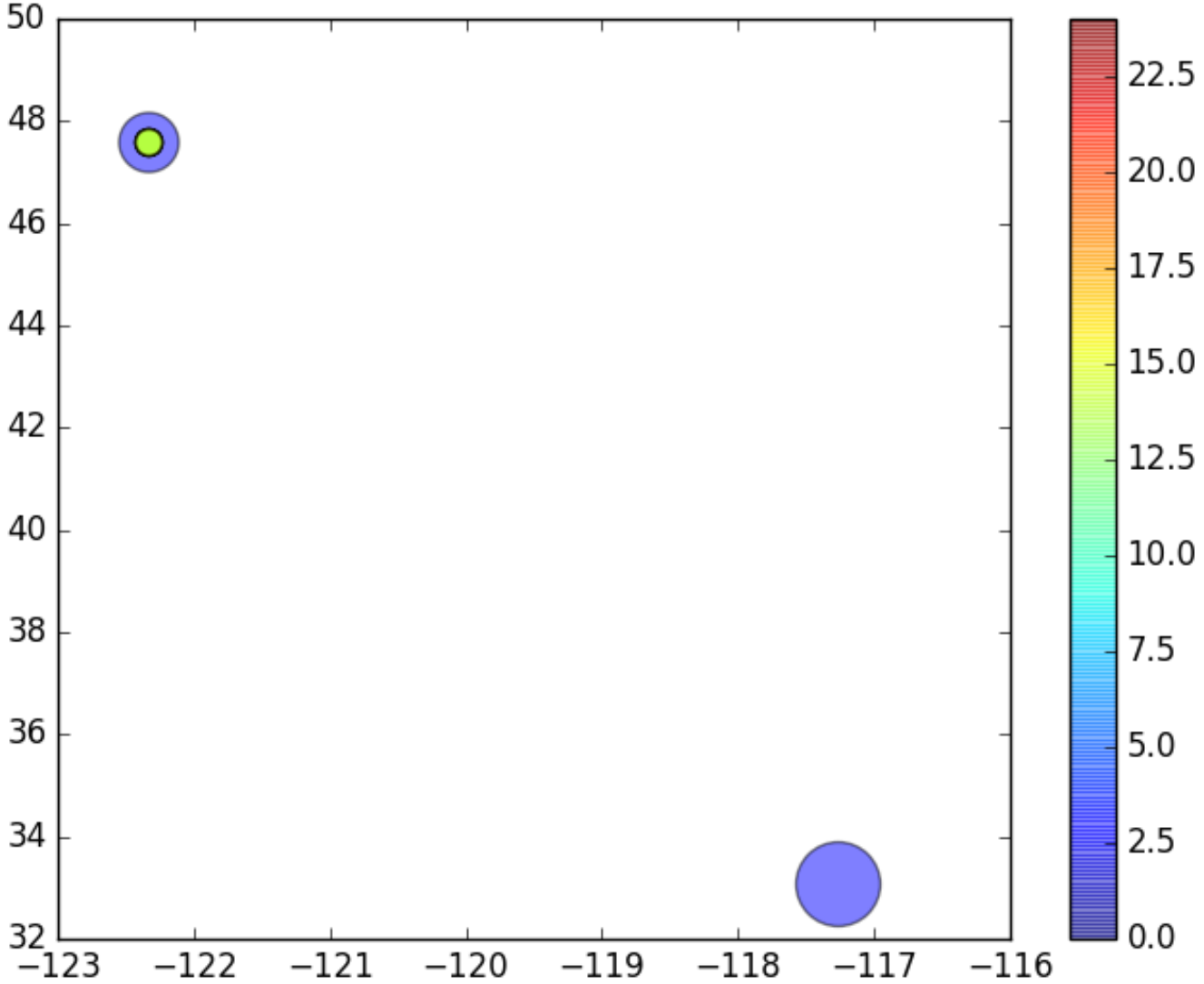
User 0 (first clustering)

[-122.34362441 47.62695596] 12
[-117.26860454 33.10603969] 340
[-117.6773782 33.58354679] 3
[-122.33783046 47.60767466] 12
[-117.23061315 32.94833509] 3
[-121.76181313 37.30174105] 2

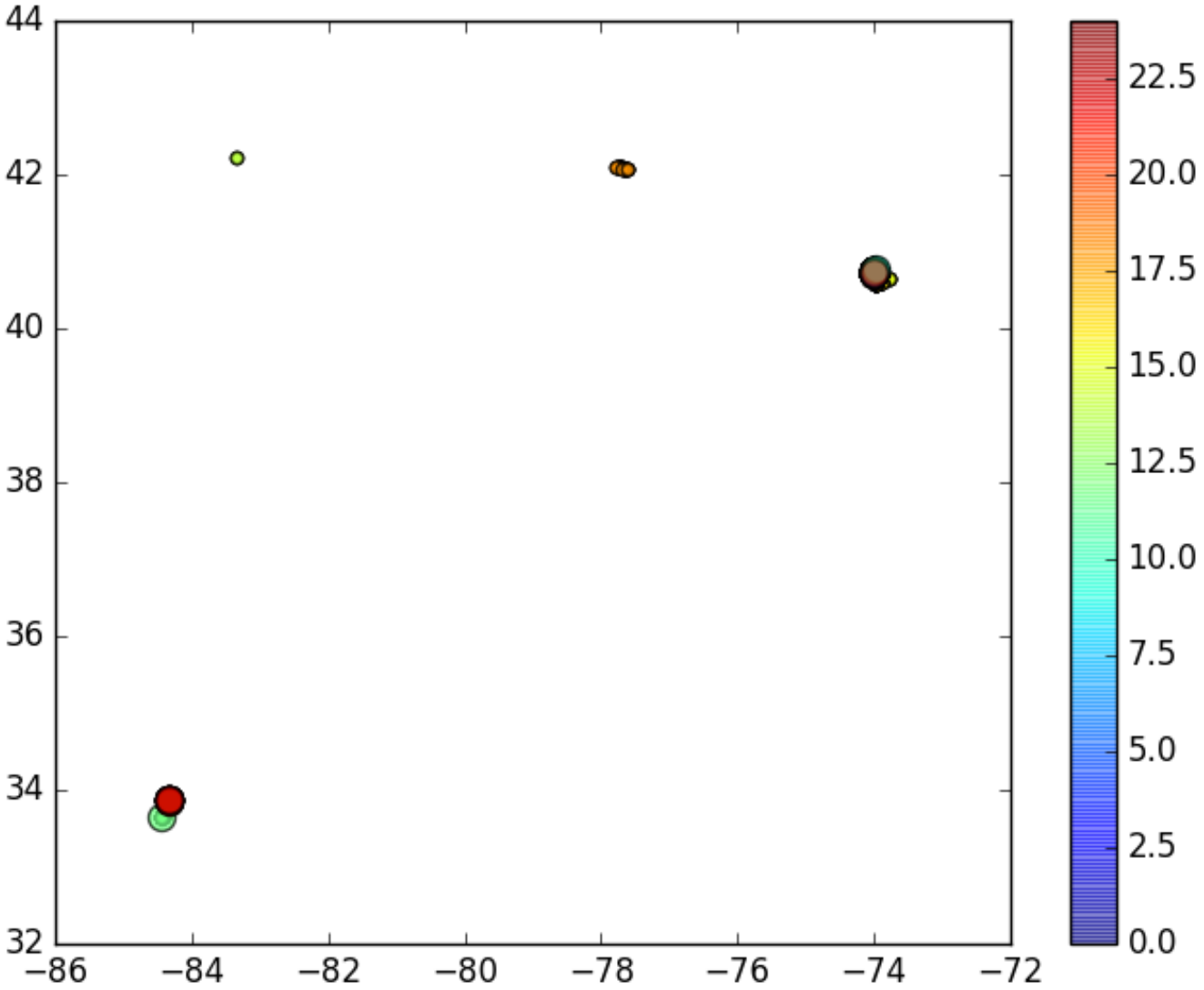
User 0 (home removed/8-18)

[-117.61704211 33.5273702] 2
[-117.7980504 33.69589998] 1
[-121.76181282 37.30174112] 1
[-117.23061315 32.94833509] 3
[-122.33692264 47.60661203] 9
[-122.34055393 47.61086255] 3

User 0


50
48
46
44
42
40
38
36
34
32
−123
−122
−121
−120
−119
−118
−117
−116
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

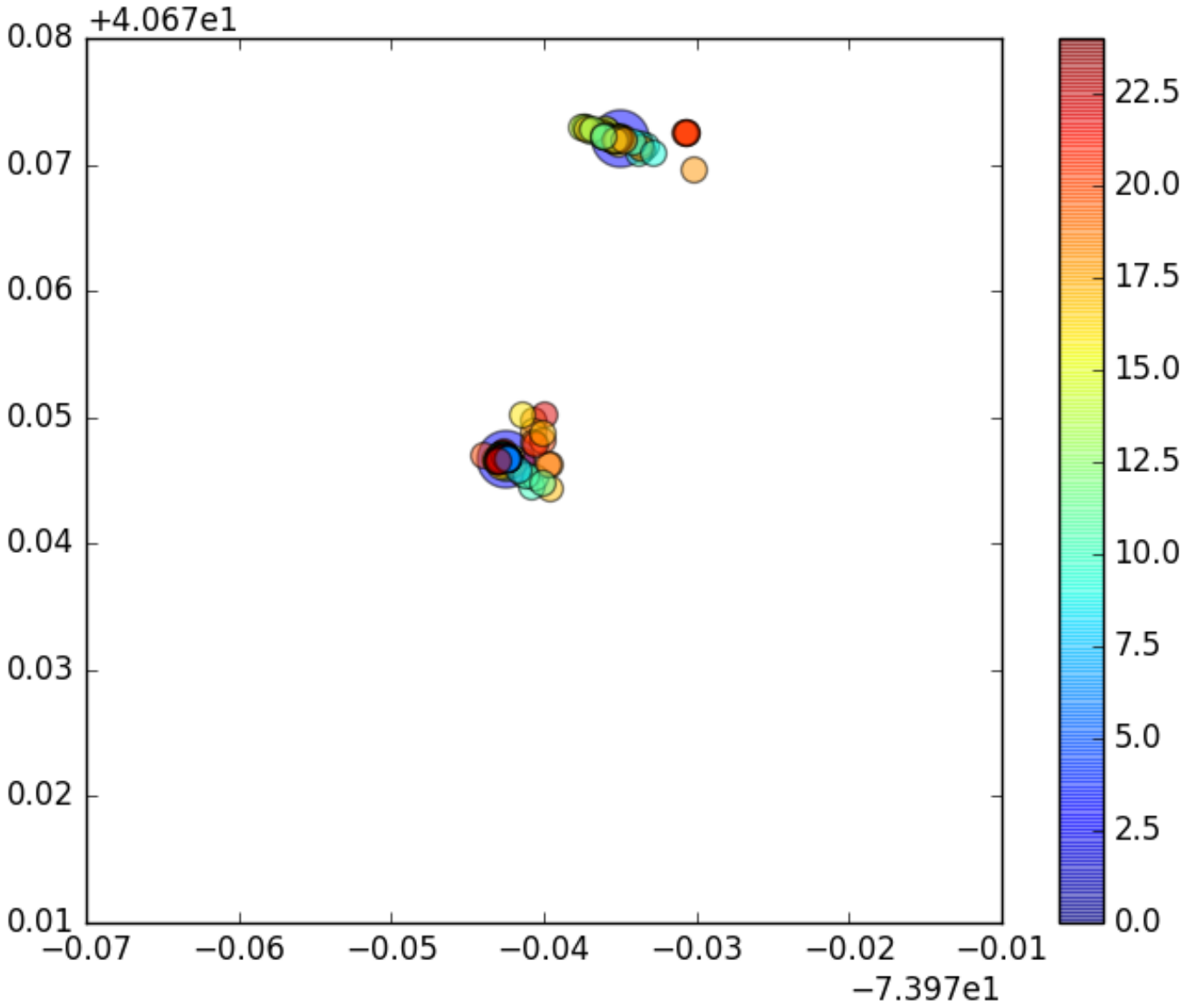
User 1 (29139375)

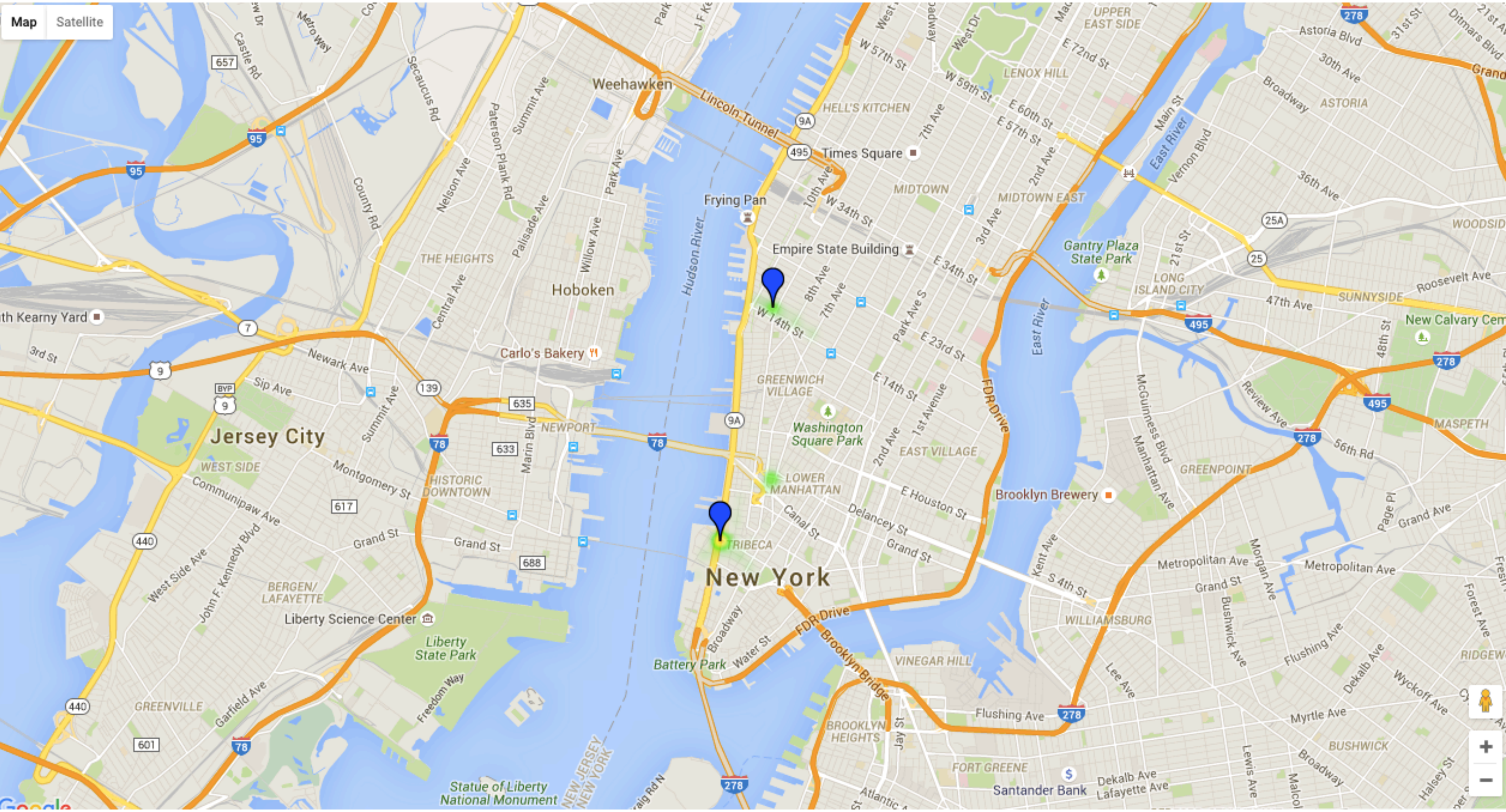


User 1

[-73.97955684 40.76579849] 2
[-74.01255563 40.71679322] 295
[-84.33022051 33.8517104] 41
[-74.00543359 40.72345232] 43
[-74.00509949 40.74214214] 81
[-74.01449044 40.71470311] 18

User 1


+4.067e1
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−0.01
−7.397e1
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0


Map
Satellite
Weehawken
Lincoln Tunnel
HELL'S KITCHEN
Times Square
MIDTOWN
MIDTOWN EAST
UPPER EAST SIDE
LENOX HILL
W 57th St
W 59th St
E 72nd St
E 60th St
E 57th St
Frying Pan
Empire State Building
W 34th St
E 34th St
W 14th St
E 23rd St
E 14th St
8th Ave
7th Ave
10th Ave
Park Ave S
1st Avenue
2nd Ave
3rd Ave
GREENWICH VILLAGE
Washington Square Park
EAST VILLAGE
LOWER MANHATTAN
TRIBECA
New York
E Houston St
Delancey St
Canal St
Grand St
FDR Drive
Brooklyn Bridge
Battery Park
Water St
Broadway
Hudson River
East River
Hoboken
THE HEIGHTS
Carlo's Bakery
NEWPORT
HISTORIC DOWNTOWN
Jersey City
WEST SIDE
BERGEN/ LAFAYETTE
Liberty Science Center
Liberty State Park
Statue of Liberty National Monument
GREENVILLE
Kearny Yard
Newark Ave
Sip Ave
Communipaw Ave
Montgomery St
West Side Ave
John F. Kennedy Blvd
Garfield Ave
Freedom Way
Summit Ave
Palisade Ave
Paterson Plank Rd
Secaucus Rd
County Rd
Nelson Ave
Central Ave
Willow Ave
Park Ave
Marin Blvd
Gantry Plaza State Park
LONG ISLAND CITY
ASTORIA
Astoria Blvd
30th Ave
36th Ave
Vernon Blvd
21st St
47th Ave
SUNNYSIDE
Roosevelt Ave
New Calvary Cem
WOODSIDE
MASPETH
56th Rd
Review Ave
GREENPOINT
McGuinness Blvd
Manhattan Ave
Brooklyn Brewery
Kent Ave
S 4th St
WILLIAMSBURG
Metropolitan Ave
Grand St
Morgan Ave
Bushwick Ave
Flushing Ave
Myrtle Ave
BUSHWICK
Broadway
Dekalb Ave
Wyckoff Ave
VINEGAR HILL
BROOKLYN HEIGHTS
FORT GREENE
Santander Bank
Dekalb Ave
Lafayette Ave
Lee Ave
Jay St
Atlantic Av
NEW JERSEY
NEW YORK
95
78
278
495
9A
25A
25
440
139
635
633
617
688
657
601
9
7

User 1 (first clustering)

[-74.01580833 40.71288009] 3

[-84.43965427 33.63974116] 1

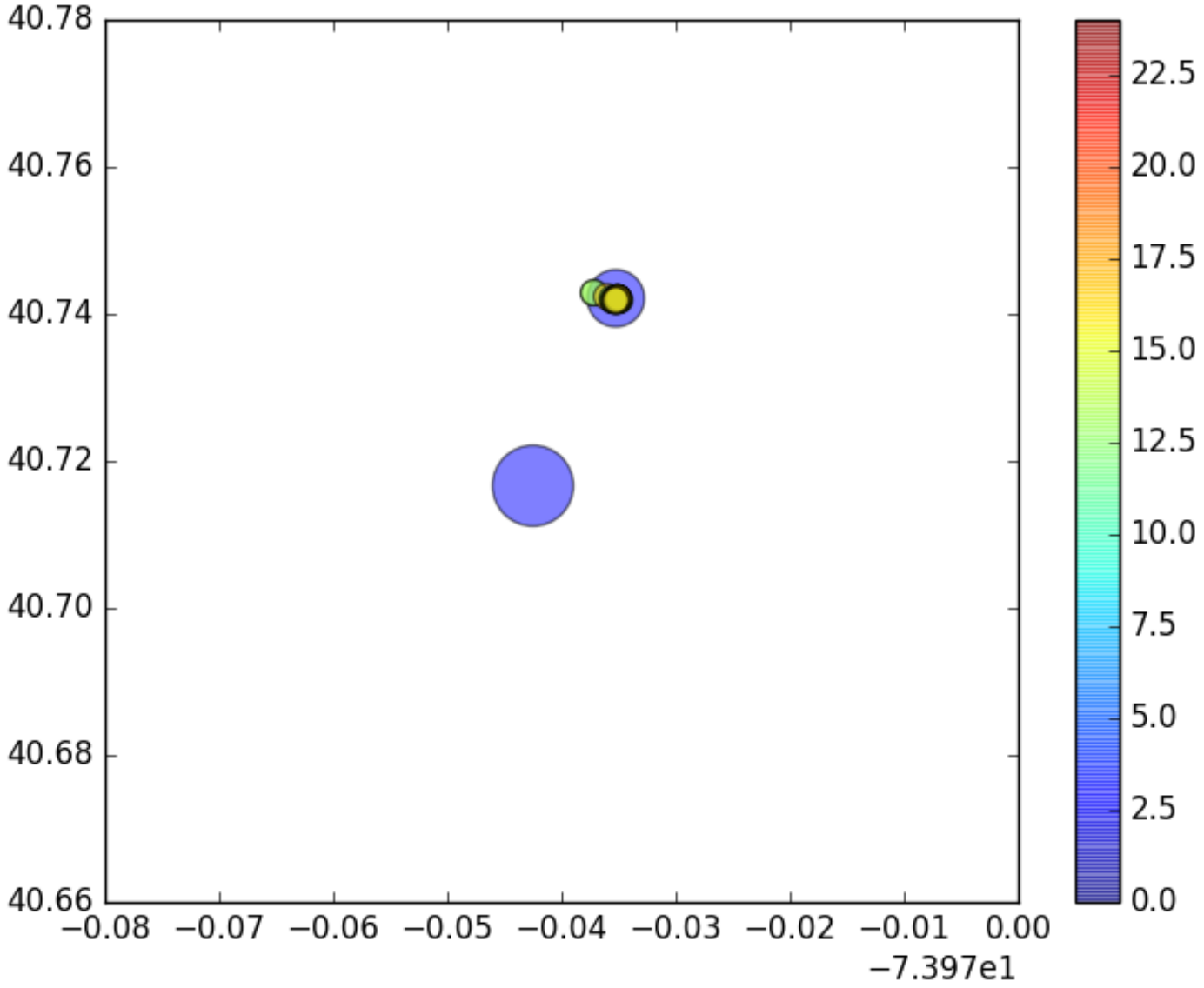
[-74.00526852 40.74215023] 47

[-84.32460583 33.86257452] 39

[-74.00528569 40.72352726] 13

[-74.0126536   40.71677521] 259 home

User 1


40.78
40.76
40.74
40.72
40.70
40.68
40.66
−0.08
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−0.01
0.00
−7.397e1
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

User 1 (home removed/8-18)

[-73.99961602 40.72680295] 1

[-74.01589685 40.71292445] 2

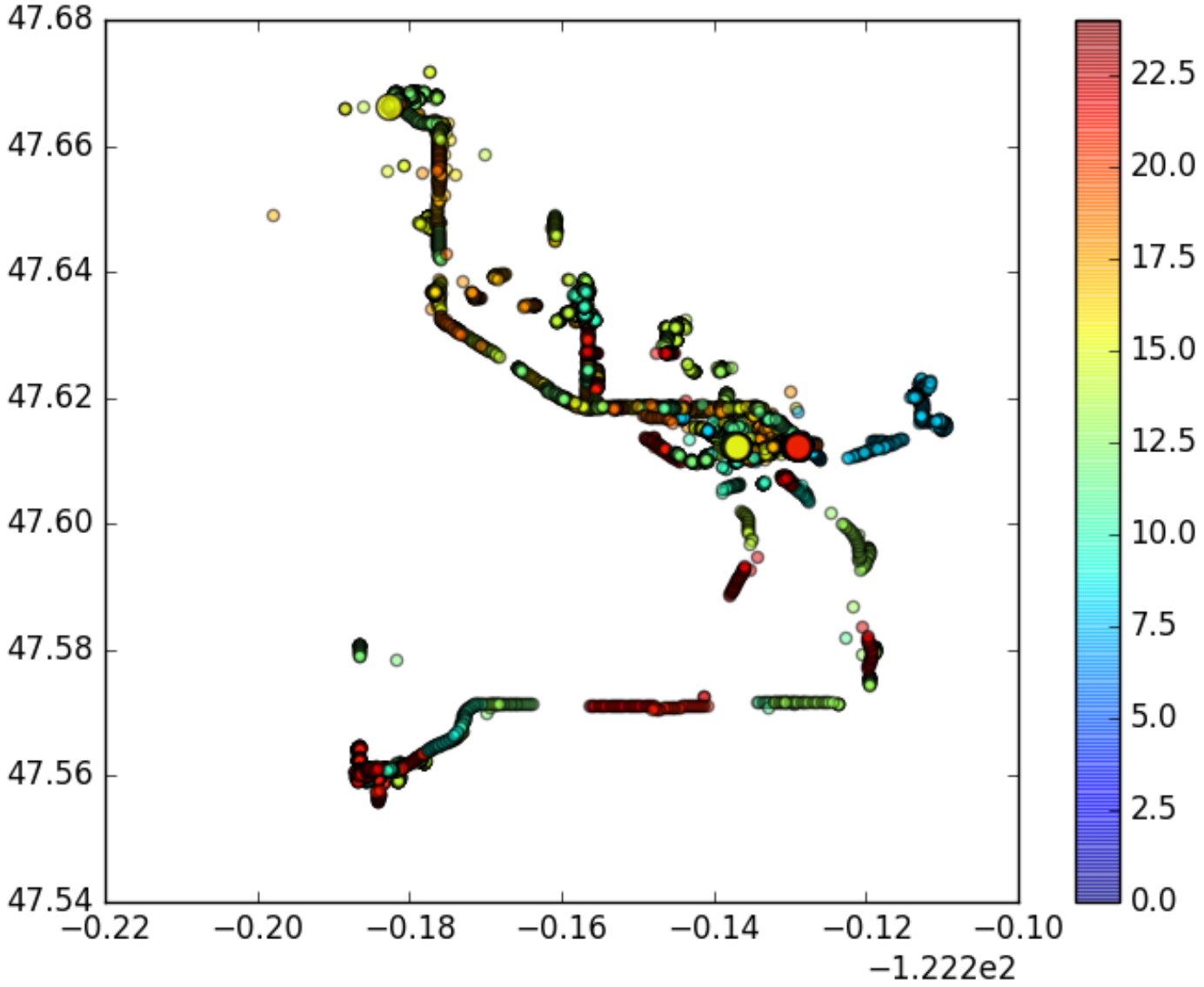
[-84.43965427 33.63974116] 1

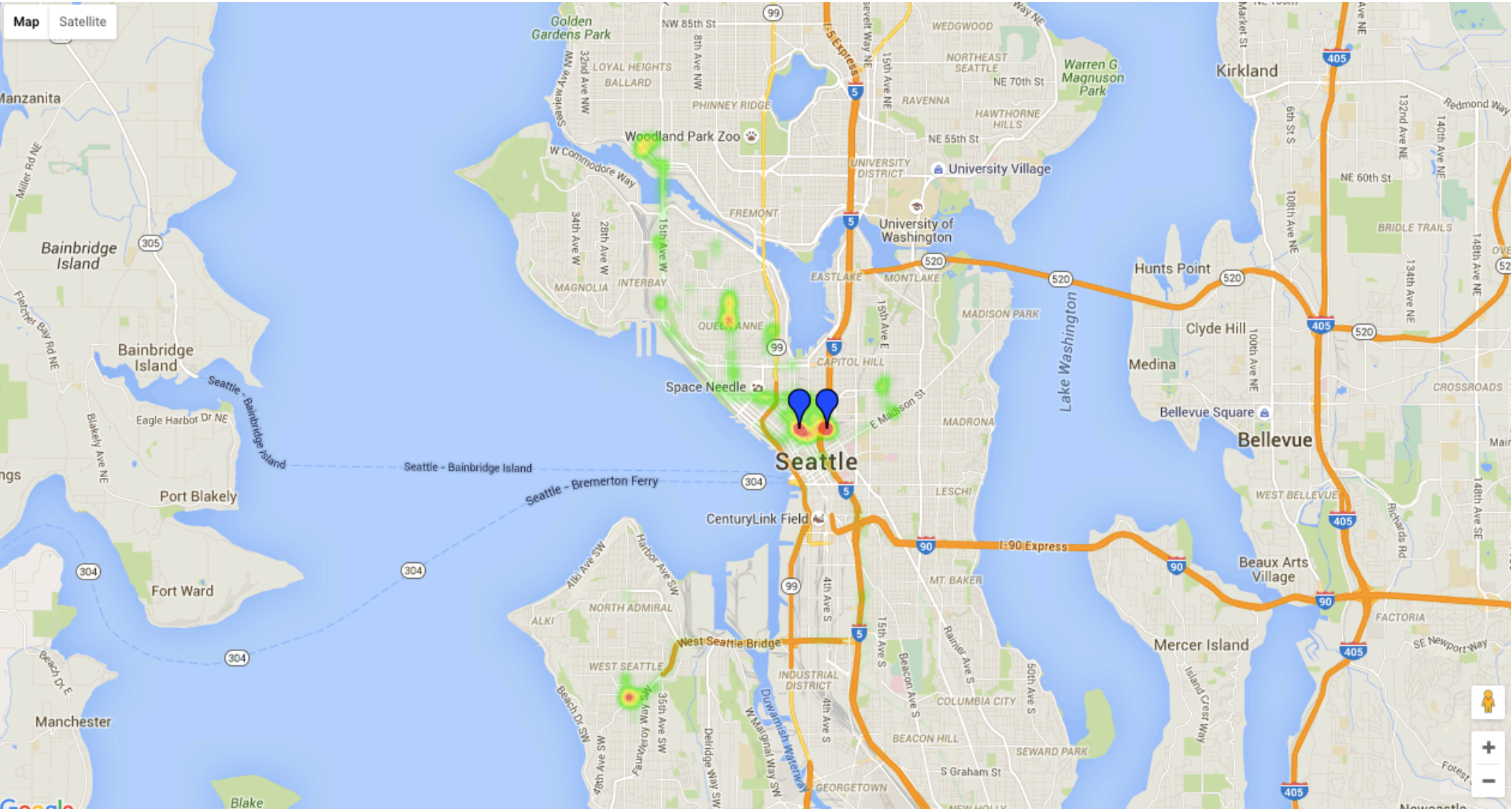
[-74.00541759 40.7234353] 11

[-74.00536778 40.7421431] 45 work

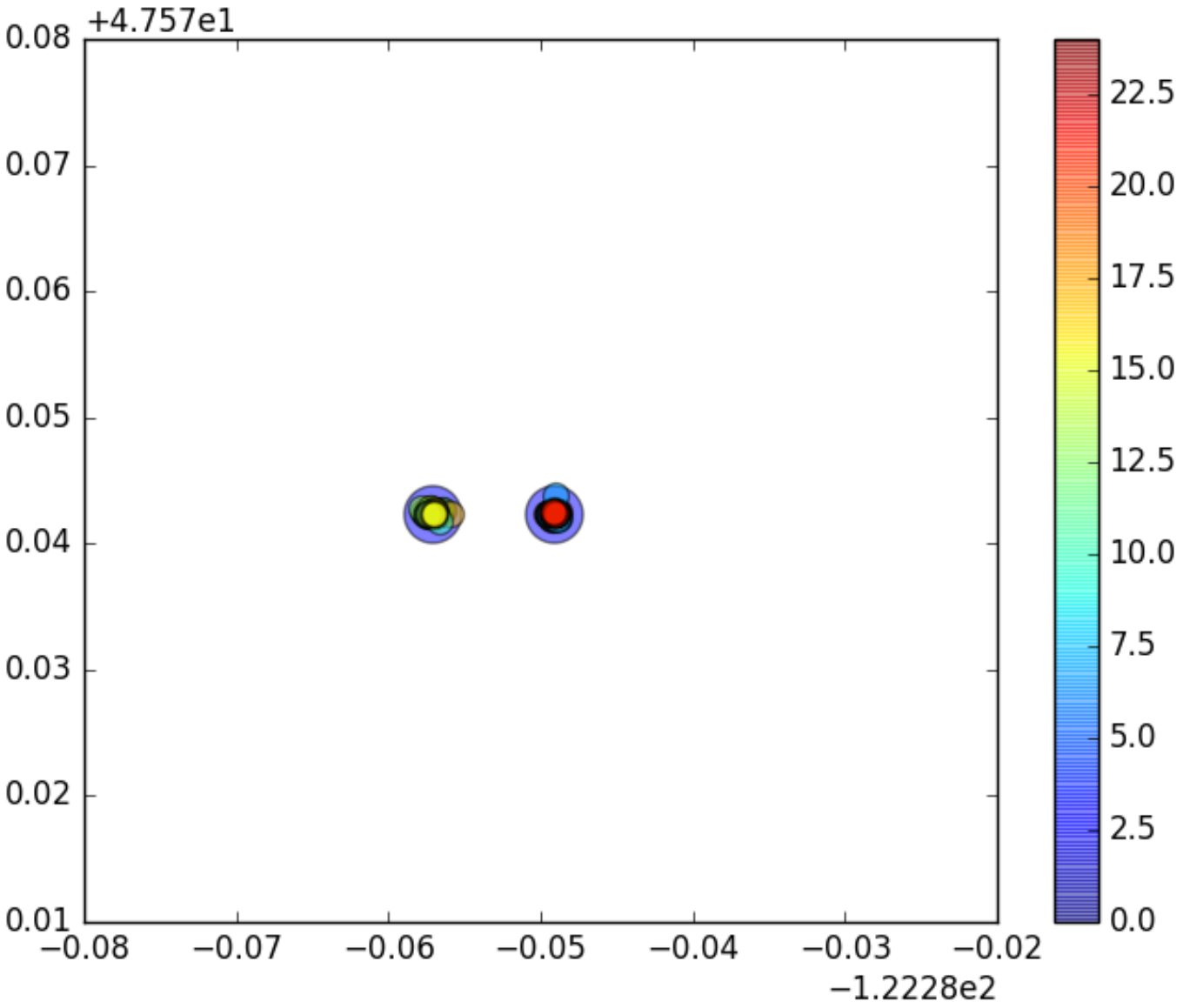
[-84.32458492 33.86254568] 13

User 2 (5935)


47.68
47.66
47.64
47.62
47.60
47.58
47.56
47.54
−0.22
−0.20
−0.18
−0.16
−0.14
−0.12
−0.10
−1.222e2
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0


Map
Satellite
Golden Gardens Park
NW 85th St
WEDGWOOD
NORTHEAST SEATTLE
NE 70th St
Warren G Magnuson Park
Kirkland
LOYAL HEIGHTS
BALLARD
PHINNEY RIDGE
RAVENNA
HAWTHORNE HILLS
Woodland Park Zoo
NE 55th St
University Village
UNIVERSITY DISTRICT
W Commodore Way
FREMONT
University of Washington
NE 60th St
BRIDLE TRAILS
Bainbridge Island
MAGNOLIA
INTERBAY
EASTLAKE
MONTLAKE
Hunts Point
MADISON PARK
Clyde Hill
QUEEN ANNE
CAPITOL HILL
Medina
Space Needle
CROSSROADS
Bellevue Square
Lake Washington
MADRONA
Seattle
Bellevue
Seattle - Bainbridge Island
Seattle - Bremerton Ferry
LESCHI
WEST BELLEVUE
CenturyLink Field
I-90 Express
Beaux Arts Village
Port Blakely
Fort Ward
MT. BAKER
NORTH ADMIRAL
ALKI
FACTORIA
West Seattle Bridge
Mercer Island
WEST SEATTLE
INDUSTRIAL DISTRICT
COLUMBIA CITY
Manchester
BEACON HILL
SEWARD PARK
S Graham St
GEORGETOWN
Duwamish Waterway
Blake

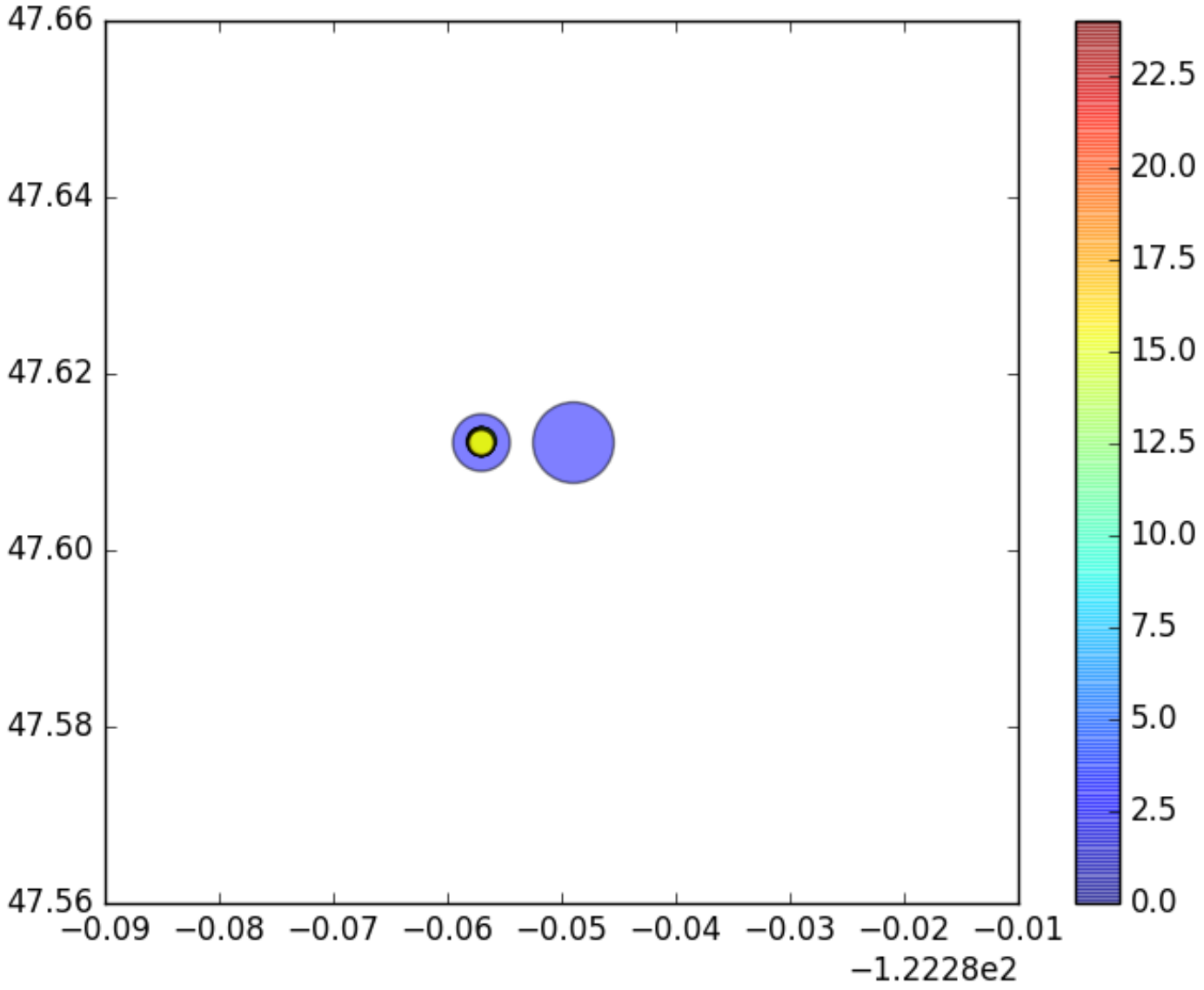
User 2


+4.757e1
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
−0.08
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−1.2228e2
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

User 2

[-122.33413085 47.61168626] 20
[-122.38165096 47.66701606] 4
[-122.33862207 47.61361091] 12
[-122.32914859 47.6123641] 264
[-122.33711099 47.612411 41] 161
[-122.33200222 47.61439294] 8

User 2


47.66
47.64
47.62
47.60
47.58
47.56
−0.09
−0.08
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−0.01
−1.2228e2
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

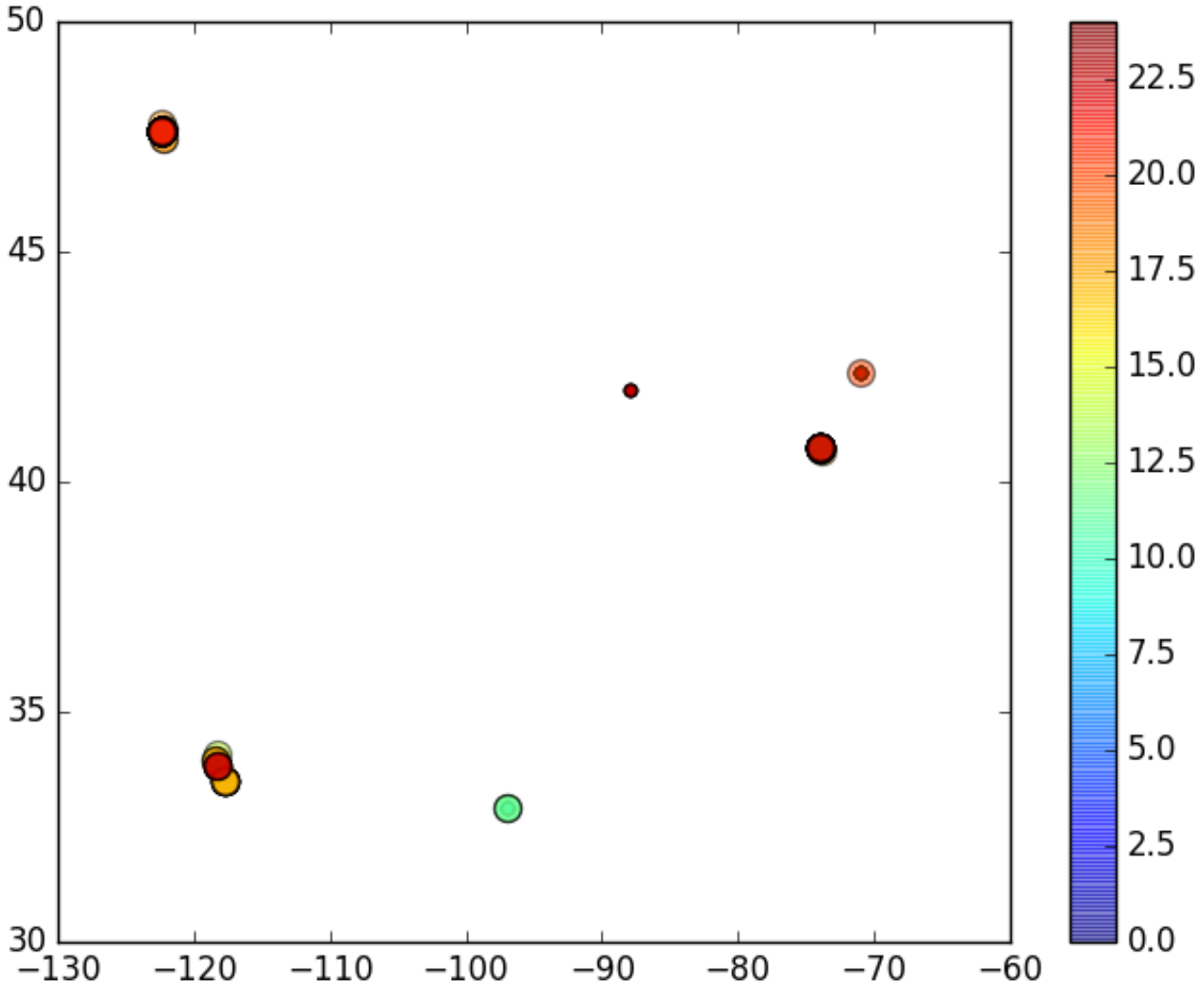
User 2 (first clustering)

- [-122.32932829 47.61233871] 64
- [-122.32908993 47.61235409] 101
  home
- [-122.33724972 47.61238483] 54
- [-122.38279714 47.66630478] 2
- [-122.32908472 47.61245667] 72
- [-122.33699465 47.61238891] 73

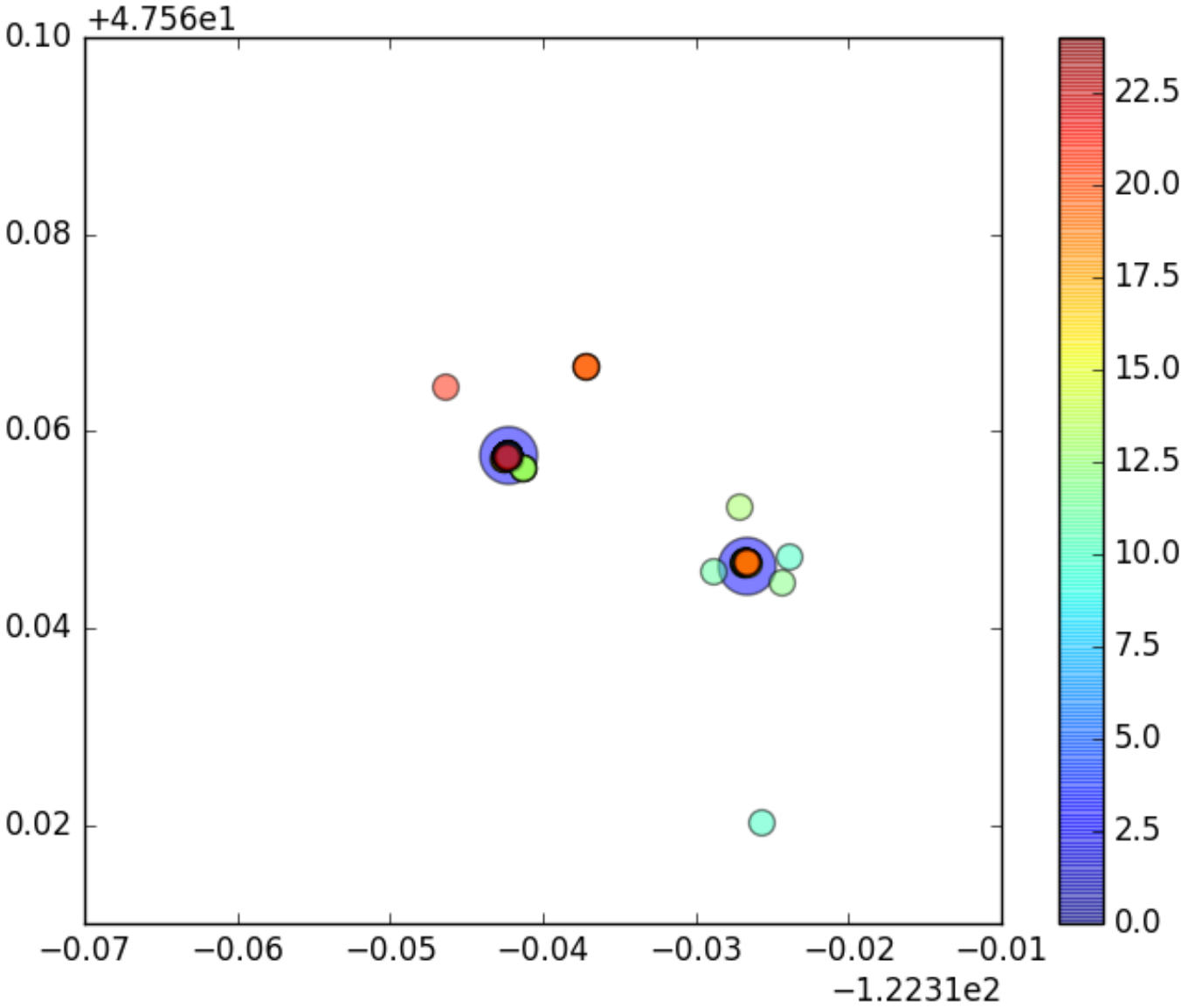
User 2 (home removed/8-18)

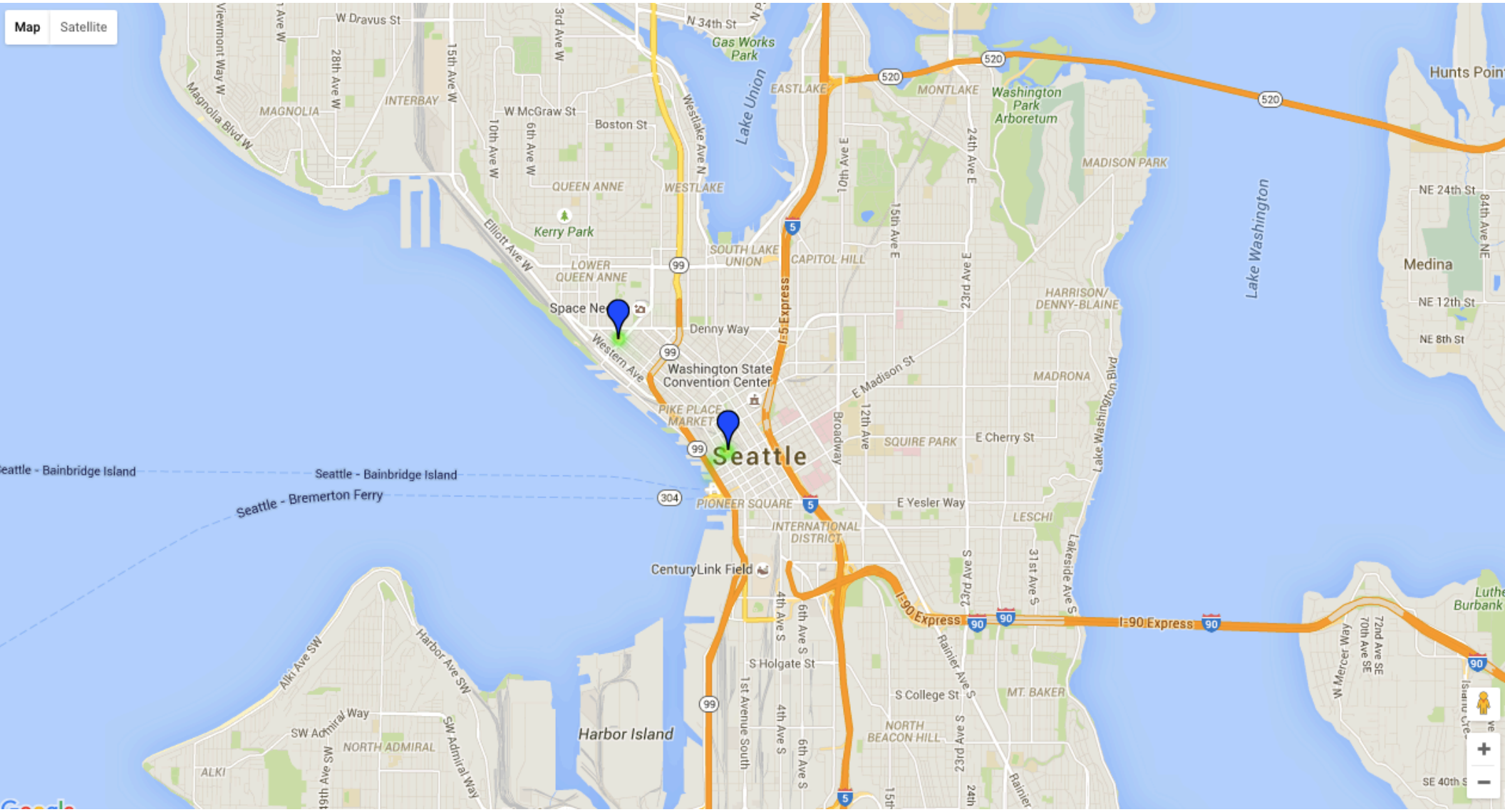
- [-122.32908795 47.61244842] 11
- [-122.33730599 47.61238051] 34
- [-122.32930327 47.6123562] 11
- [-122.33691781 47.61238139] 34
- [-122.33709298 47.61239436] 59 work
- [-122.38279714 47.66630478] 2

User 3 (71840)


50
45
40
35
30
−130
−120
−110
−100
−90
−80
−70
−60
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

User 3

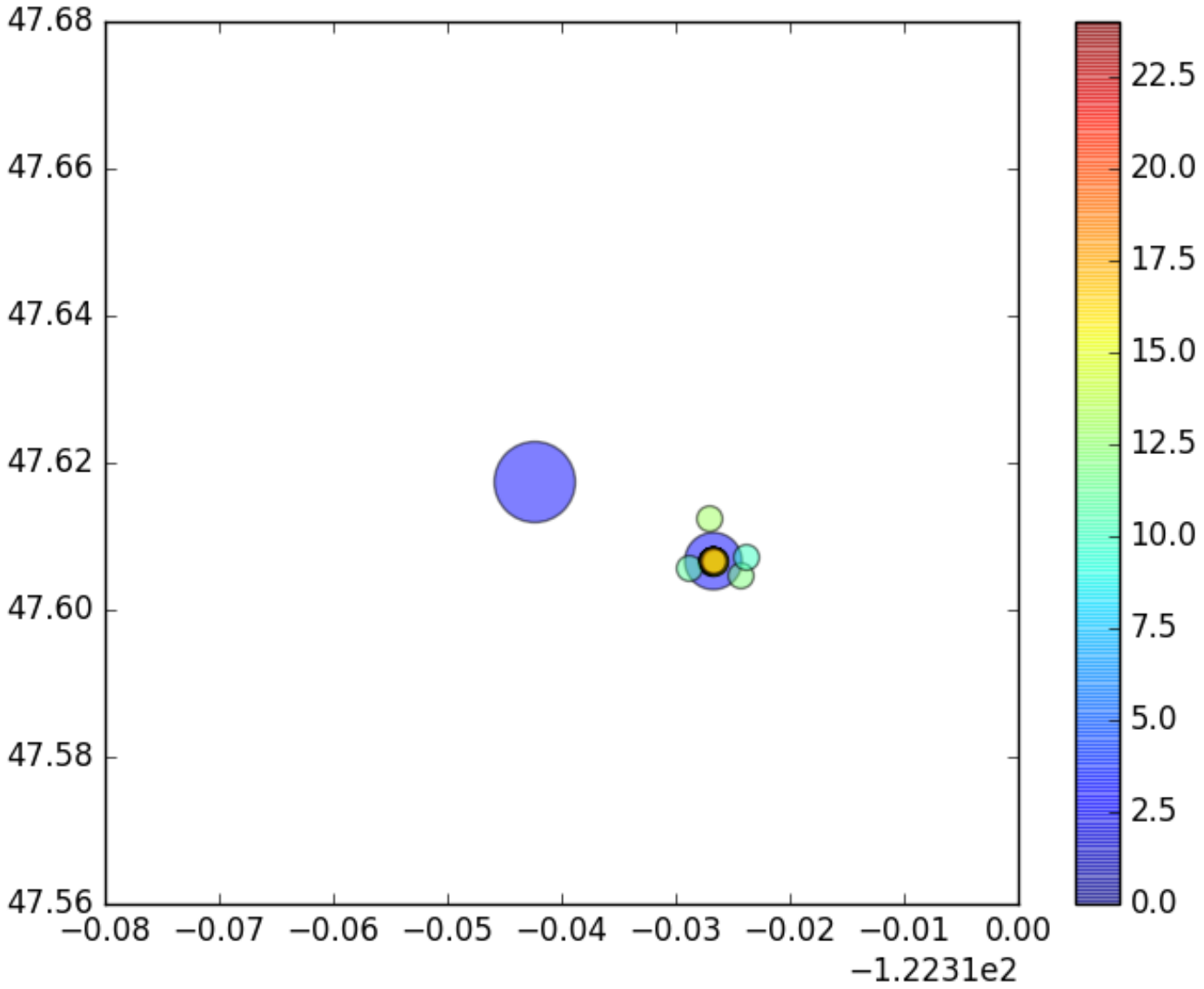

+4.756e1
0.10
0.08
0.06
0.04
0.02
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−0.01
−1.2231e2
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0


Map
Satellite
Seattle
Space Needle
Kerry Park
Gas Works Park
Washington Park Arboretum
Washington State Convention Center
CenturyLink Field
Harbor Island
Lake Union
Lake Washington
Seattle - Bainbridge Island
Seattle - Bremerton Ferry
I-5 Express
I-90 Express
MAGNOLIA
INTERBAY
QUEEN ANNE
WESTLAKE
EASTLAKE
MONTLAKE
MADISON PARK
SOUTH LAKE UNION
CAPITOL HILL
HARRISON/ DENNY-BLAINE
MADRONA
SQUIRE PARK
PIKE PLACE MARKET
PIONEER SQUARE
INTERNATIONAL DISTRICT
LESCHI
MT. BAKER
NORTH BEACON HILL
NORTH ADMIRAL
ALKI
Medina
Hunts Point
W Dravus St
N 34th St
W McGraw St
Boston St
Denny Way
E Madison St
E Cherry St
E Yesler Way
S Holgate St
S College St
SW Admiral Way
Elliott Ave W
Western Ave
Broadway
Lakeside Ave S
Lake Washington Blvd
NE 24th St
NE 12th St
NE 8th St
SE 40th St

User 3

- [-122.35229402 47.61768209] 209 home
- [-122.33669855 47.60636679] 67 work
- [-73.99368704 40.74981457] 30
- [-117.96986162 33.63792486] 29
- [-71.01341222 42.35855098] 1
- [-97.02951421 32.90100695] 3

User 3


47.68
47.66
47.64
47.62
47.60
47.58
47.56
−0.08
−0.07
−0.06
−0.05
−0.04
−0.03
−0.02
−0.01
0.00
−1.2231e2
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0

User 3 (first clustering)

- [-122.35234878 47.61751826] 205 home
- [-117.96986162 33.63792486] 29
- [-97.02951421 32.90100695] 3
- [-122.33669855 47.60636679] 67
- [-73.89754914 40.8017093] 31
- [-122.34948789 47.62607793] 4

User 3 (home removed/8-18)

- [-73.99198955 40.74957177] 7
- [-122.34715979 47.62657464] 1
- [-117.81673243 33.56950591] 10
- [-122.33572613 47.58035328] 1
- [-122.33670952 47.606766618] 59 work
- [-97.02951421 32.90100695] 3

Results

- userid 12962060 home [-117.26860454 33.10603969] work [-122.33783046 47.60767466]
- userid 29139375 home [-74.0126536 40.71677521] work [-74.00536778 40.7421431]
- userid 5935 home [-122.32909364 47.61235342] work [-122.33709602 47.61239253]
- userid 71840 home [-122.35229402 47.61768209] work [-122.33670168 47.60666858]

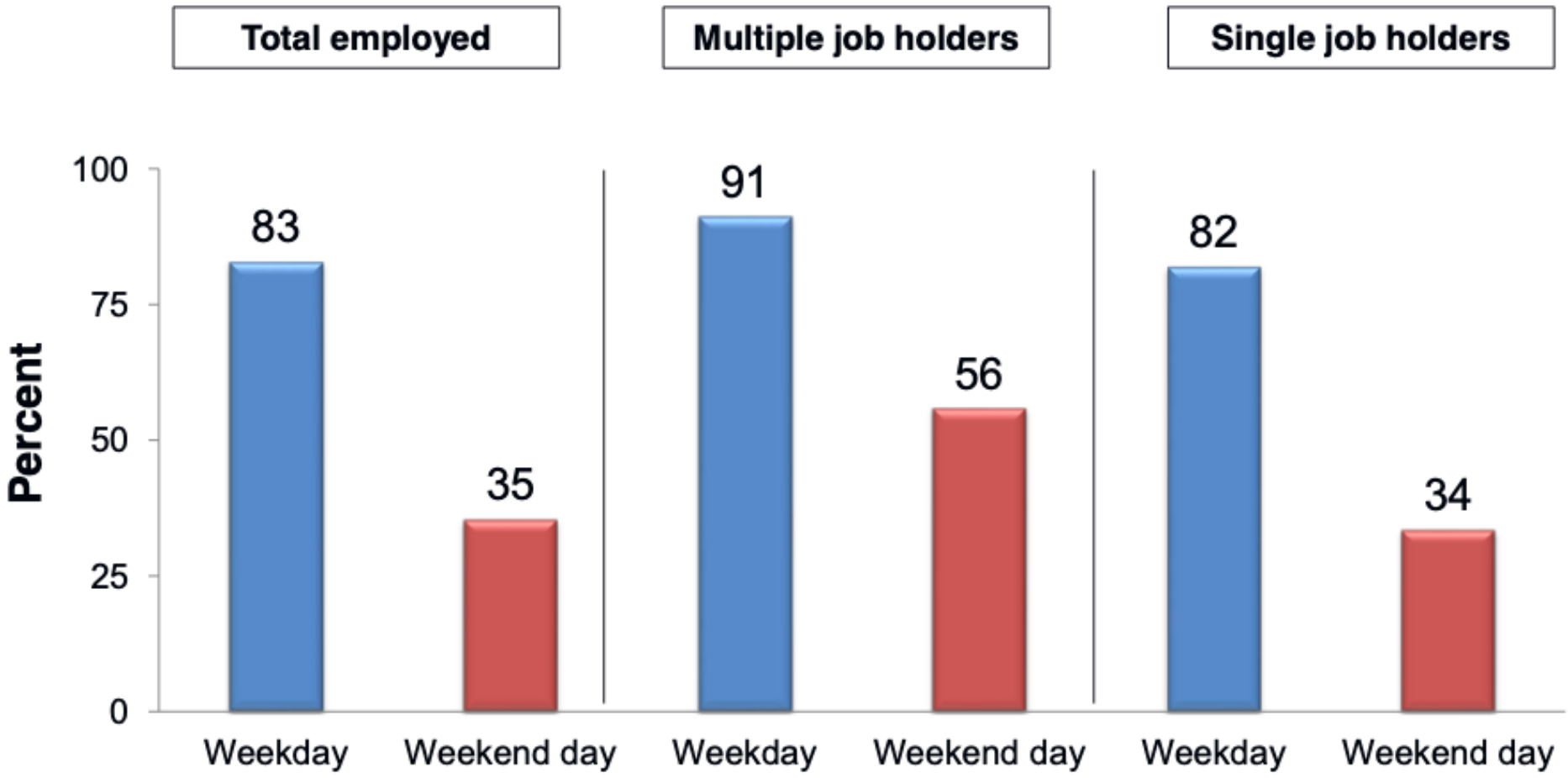
Discussion

- I assumed users work one job and mostly during the day time.
- For night shift, we can remove the restriction for local time in the code and the work location will be identified as the location a user spend second most of his time at.
- We can modify the code to predict more than one work locations (preserve more centers).
- Identify unemployed people from the data? Difference with work from home?

Future work

- Local time dependence of locations
- Weekend
- Time gaps (identical locations on both ends)
- Sequence of locations (from home to work on weekday mornings)
- Train the model with known home/work locations (# centers, hvar threshold, local time restriction for work)

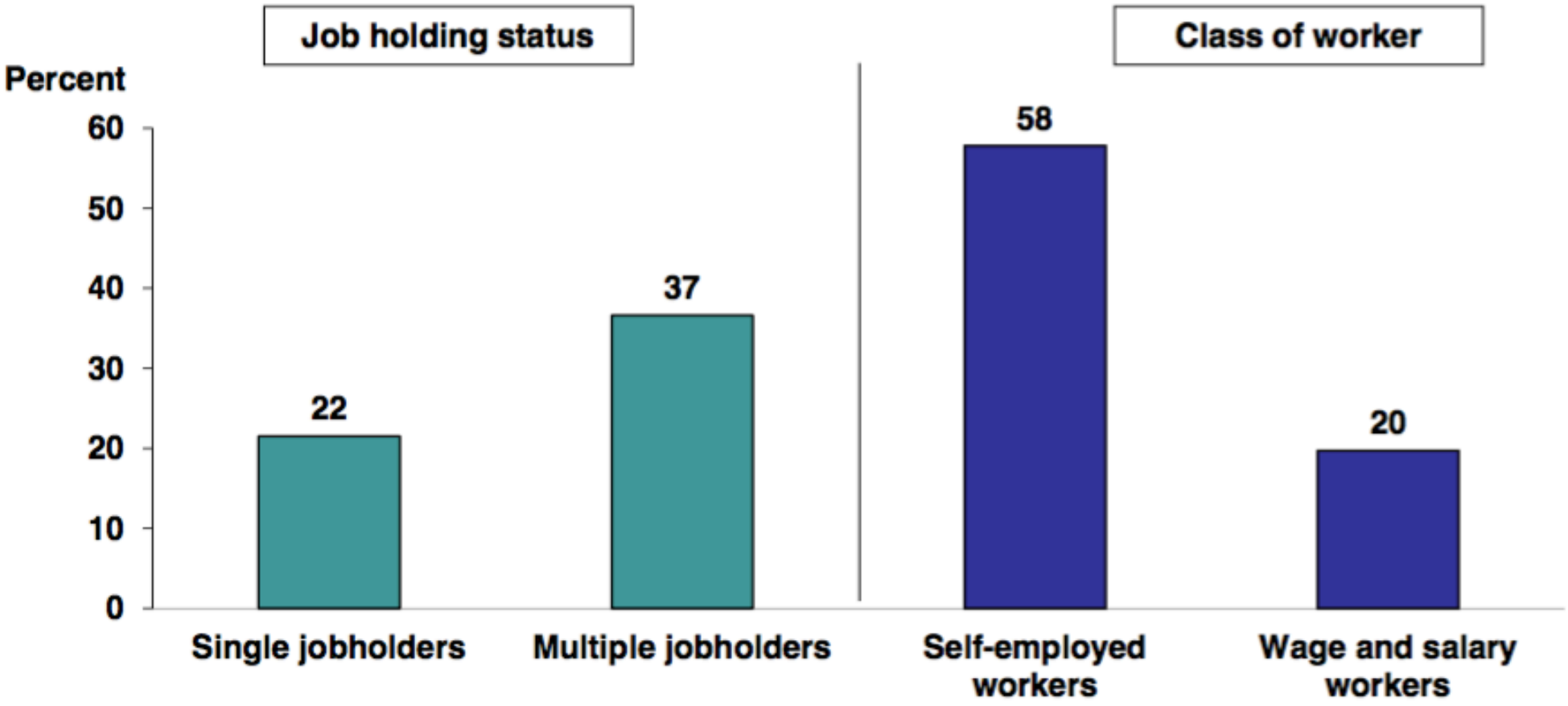
Percent of population who worked on weekdays and weekend days



NOTE: Data include all persons age 15 and over. Weekdays are defined as non-holiday weekdays. Holidays are included with weekend days. Data are annual averages for 2014.

SOURCE: Bureau of Labor Statistics, American Time Use Survey

Percent of employed persons who worked at home on an average workday



NOTE: Data include all employed persons age 15 and over on days they worked. Data include all days of the week and are annual averages for 2014.

SOURCE: Bureau of Labor Statistics