Driving Distances and Travel Times Using SAS and Google Maps

Presentation to Institutional Research Council

CUNY, New York

August 26, 2011

Presenter:

Anatoliy Kadinsky, Associate Director for Research

Hunter College,CUNY

Session Objectives:

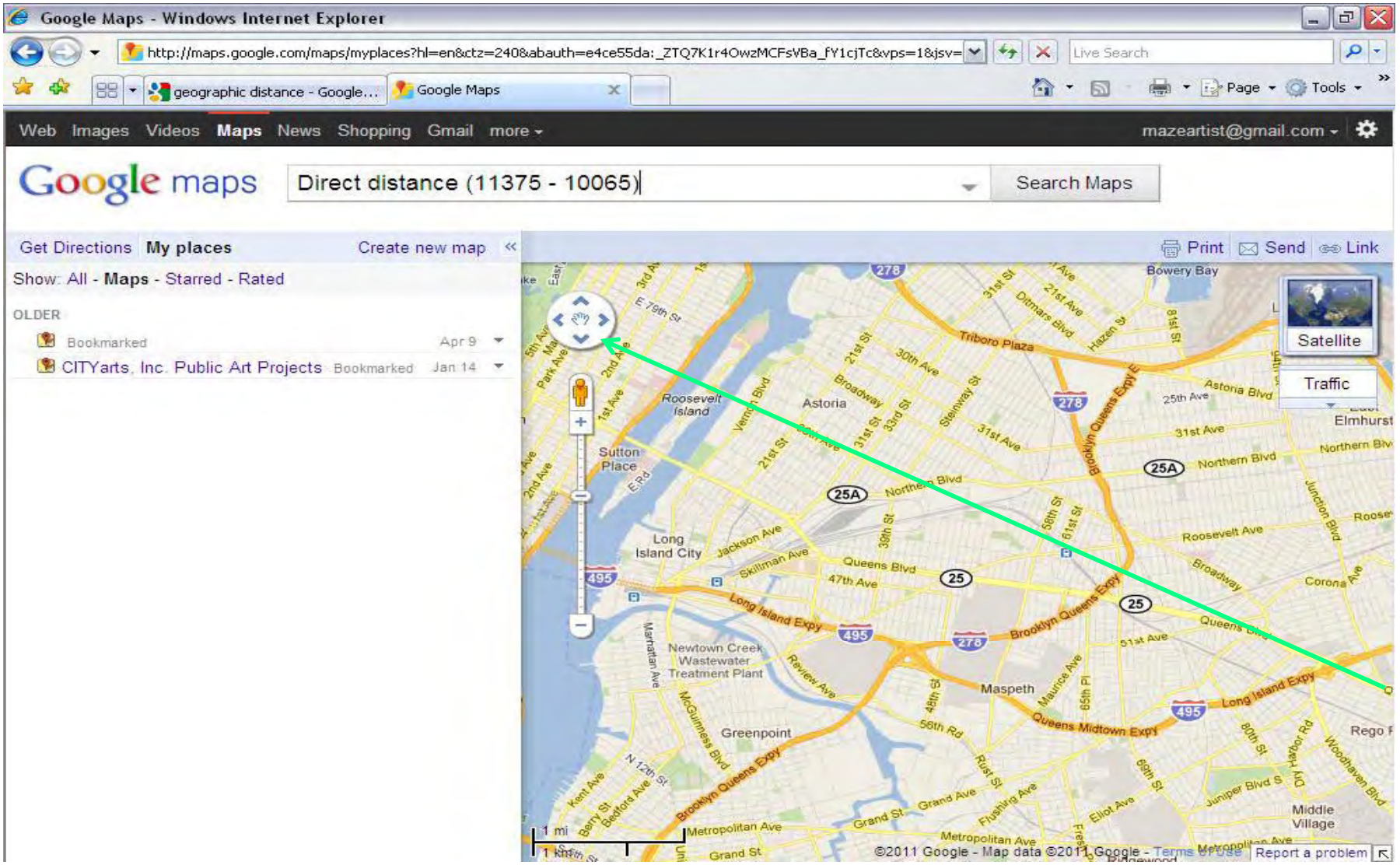
- Describe importance and usage of travel time information for the research purposes
- Describe strategies for accomplishing distance and travel time calculation
- Discuss pros and cons of using SAS and Google Maps for distances and travel times calculations

Information Available from using SAS and Google Maps

- Straight line distance (geographic distance)

- Driving distance
- Driving time
- Public Transit travel time
- Walking distance and time
- Bicycling distance and time

Straight Line Distance



Straight Line Distance

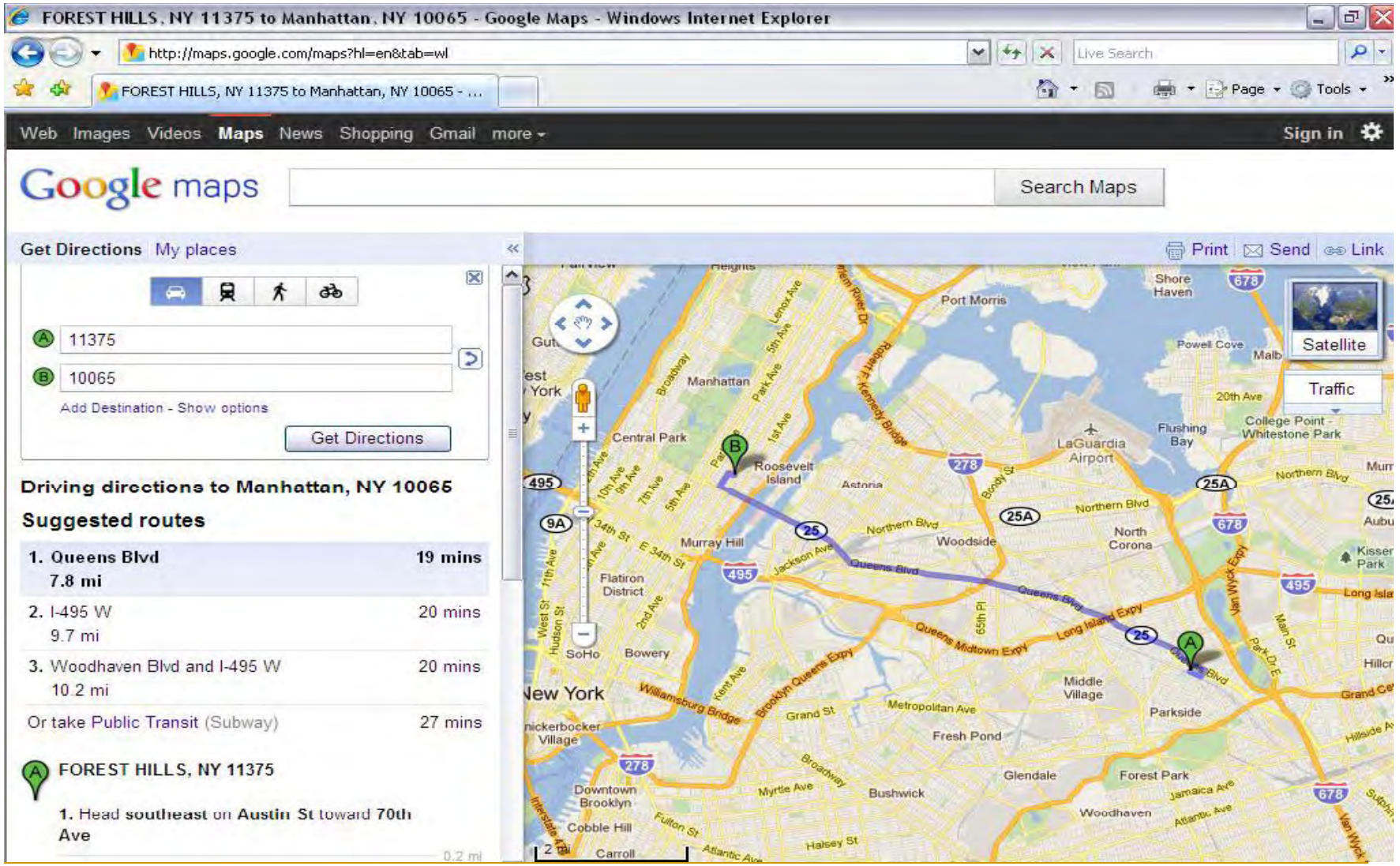
Input: Zip codes of origin and destination

- Geocodes from SASHelp zipcodes dataset
- Haversine formula calculates great-circle distances between 2 points on a sphere from their longitudes and latitudes

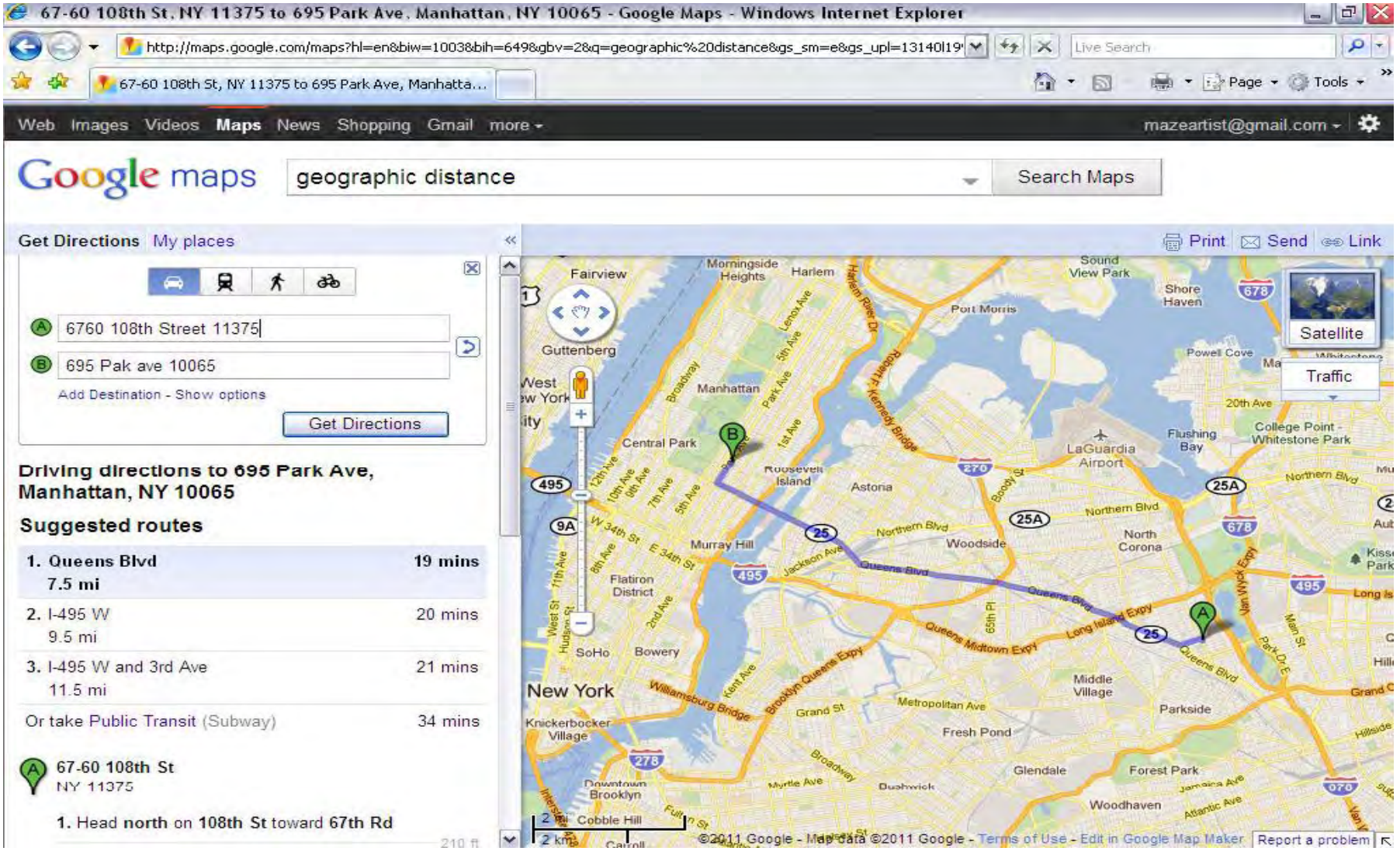
G_distance = 3949.99 * arcos(sin(lat1) * sin(lat2) + cos(lat1) * cos(lat2) * cos(long2 – long1));

Output: straight line distance between centroids of 2 zips

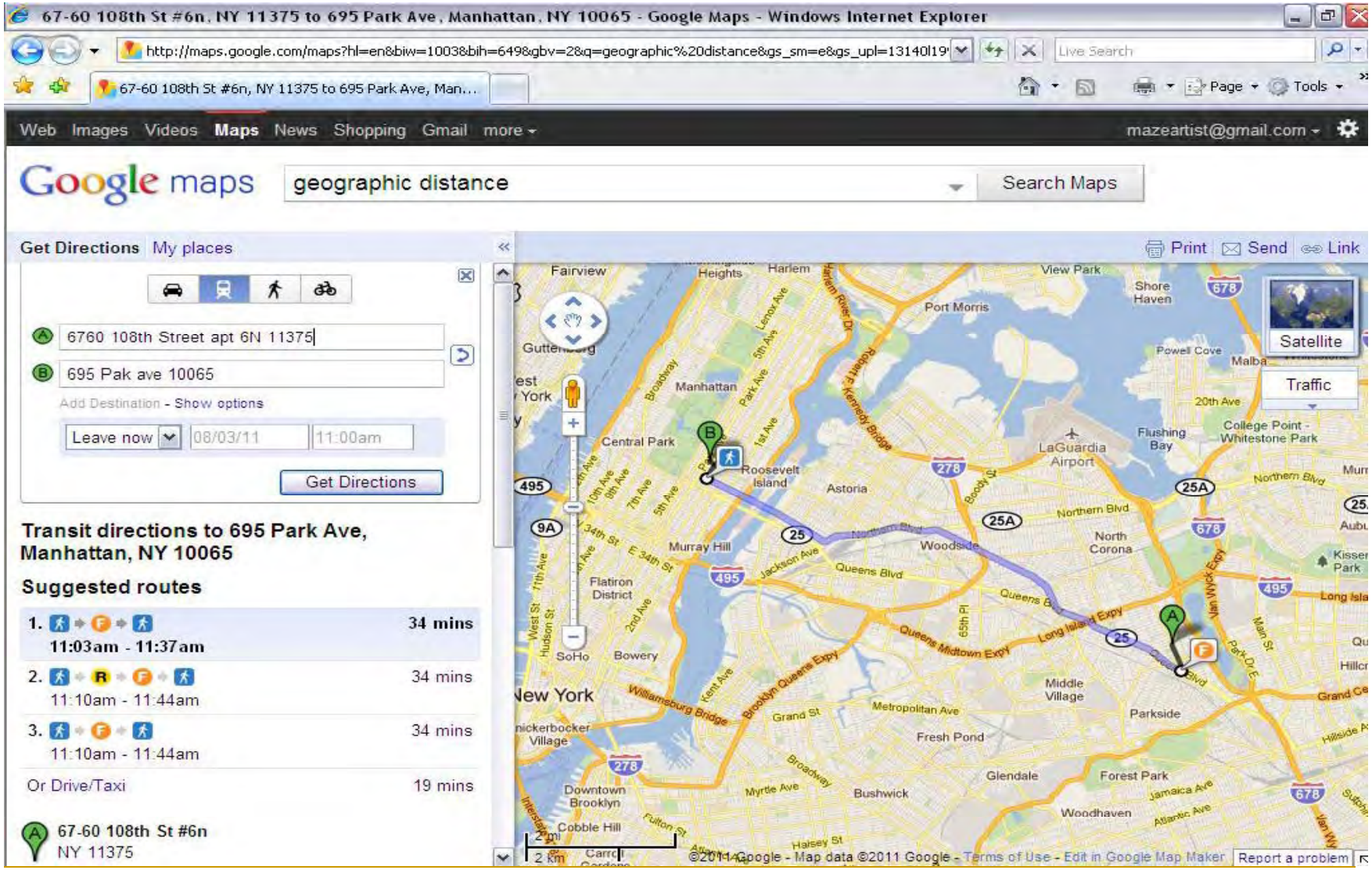
Driving Distance and Time (by zipcode)



Driving Distance and Time (by address)



Public Transit Travel Time



Basic Steps to Capture Distance/Time

- Prepare input file which consists of pairs of zips, addresses or lat/long coordinates
- Run a MACRO in a SAS program, which creates FILENAME statement
- FILENAME statement takes a pair of input parameters, places them into URL to access Google Maps and returns a webpage source
- Locate distance/time data in the page source file (html)
- Parse out distance/time data from html
- Append a new record to the base containing prior address distance/time record

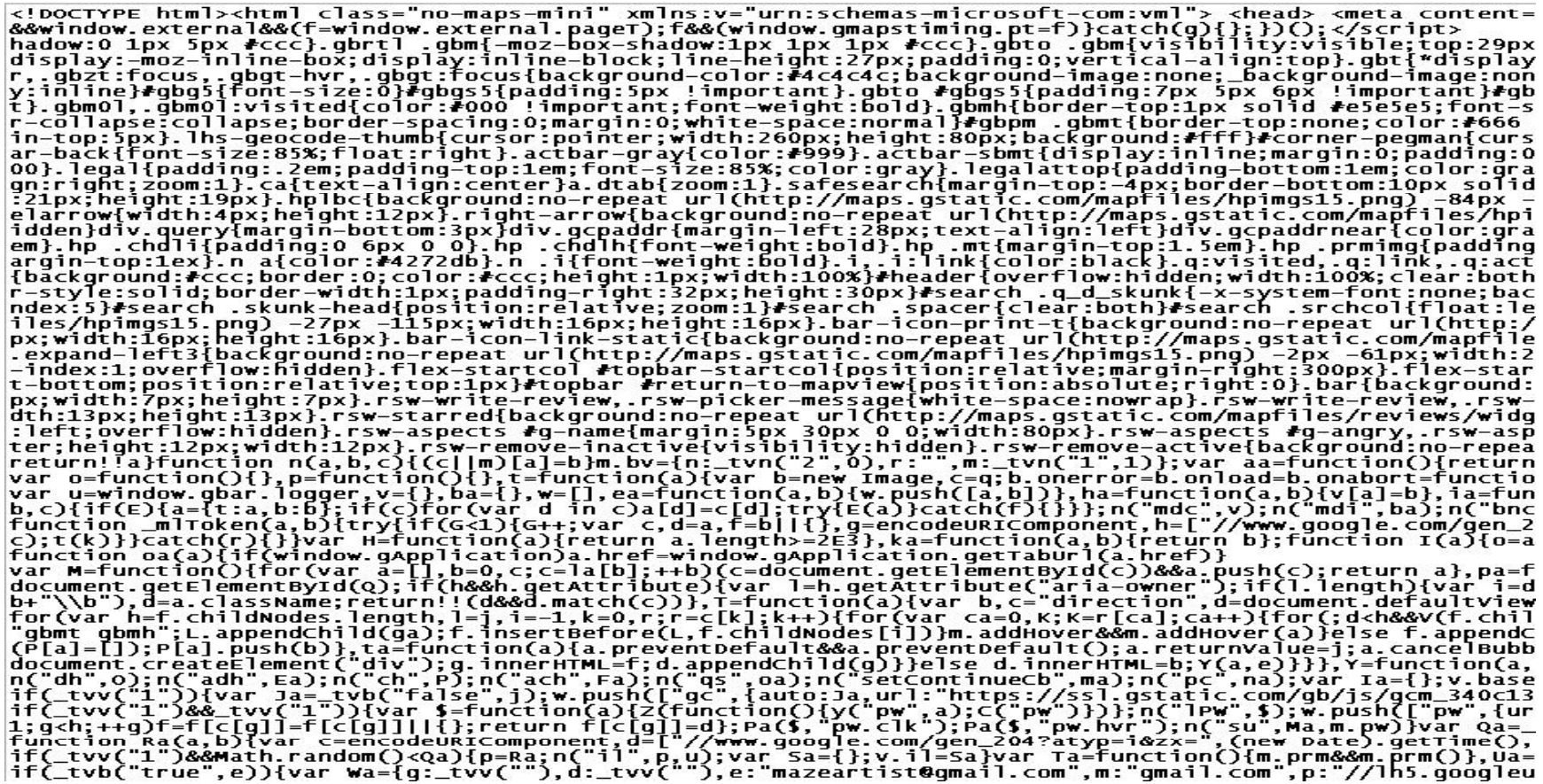
Basic Steps to Capture Distance/Time

- Run a MACRO in a SAS program, which creates FILENAME statement

```
* place number of addresses in a macro variable;
data _null_;
call symputx('naddr',obs);
stop;
set addresses nobs=obs;
run;

* use a loop within a macro to access Google Maps multiple time;
%macro distance_time(type);
.............................

%do j=1 %to &naddr;
data _null_;
nrec = &j;
set addresses point=nrec;
call symputx('id',id);
call symputx('a1',translate(trim(addr1),'+',' '));
call symputx('a2',translate(trim(addr2),'+',' '));
stop;
run;

filename x url "http://maps.google.com/maps?daddr=&a1=%nrstr(&saddr)=&a2%nrstr(&dirflg)=&type";
filename z temp;
.............................................
.............................................
```

Basic Steps to Capture Distance/Time

- Web page source code (HTML)



Basic Steps to Capture Distance/Time

- Locate distance/time data in the page source file (html)
- Parse out distance/time data from html

```
infile z recfm=f lrecl=&filesize. eof=done;

%if &type ^= tr %then %do;
input @ "distance:'" @;
input text $50.;
distance = input(scan(text,1,"' "),comma12.);
units    = scan(text,2,"' ");
text     = scan(text,3,"'");
%end;
%else %do;
input @ "Travel time: about" @;
input text $50.;
text = scan(text,2,'<>');
%end;

select;
 when (find(text,'day') ne 0)
  time = 86400*input(scan(text,1,' '),best.)+3600*input(scan(text,3,' '),best.);
 when (find(text,'hour') ne 0)
  time = 3600*input(scan(text,1,' '),best.)+60*input(scan(text,3,' '),best.);
 otherwise
  time = 60*input(scan(text,1,' '),best.);
end;
 tr_time=time/60;
output;
```

SAS Program Output

VIEWTABLE: Work.Distance_time

	ADDRESS #1	ADDRESS #2	ID #	TIME (MIN)
16	8315 263RD ST 11004	65-30 KISSENA BLVD	.	49
17	8321 258TH ST 11004	65-30 KISSENA BLVD	.	48
18	8033 255TH ST 11004	65-30 KISSENA BLVD	.	43
19	7810 266TH ST 11004	65-30 KISSENA BLVD	.	42
20	8003 256TH ST 11004	65-30 KISSENA BLVD	.	38
21	8020 266TH ST 11004	65-30 KISSENA BLVD	.	45
22	2505 40TH AVE 11101	65-30 KISSENA BLVD	.	54
23	1136 44TH DR 11101	65-30 KISSENA BLVD	.	51
24	2818 41ST AVE 11101	65-30 KISSENA BLVD	.	51
25	1151 44TH RD 11101	65-30 KISSENA BLVD	.	51
26	1017 47TH AVE 11101	65-30 KISSENA BLVD	.	57
27	3414 43RD ST 11101	65-30 KISSENA BLVD	.	47
28	4444 21ST ST 11101	65-30 KISSENA BLVD	.	50
29	3422 44TH ST 11101	65-30 KISSENA BLVD	.	49
30	3405 44TH ST 11101	65-30 KISSENA BLVD	.	48
31	3607 STEINWAY ST 11101	65-30 KISSENA BLVD	.	49
32	3408 43RD ST 11101	65-30 KISSENA BLVD	.	47
33	PO BOX 1205 11101	65-30 KISSENA BLVD	.	57
34	3721 GREENPOINT AVE 11101	65-30 KISSENA BLVD	.	55
35	1124 46TH AVE 11101	65-30 KISSENA BLVD	.	54
36	3421 37TH ST 11101	65-30 KISSENA BLVD	.	47
37	PO BOX 1594 11101	65-30 KISSENA BLVD	.	57
38	1076 JACKSON AVE 11101	65-30 KISSENA BLVD	.	61
39	1112 47TH AVE. APT 11101	65-30 KISSENA BLVD	.	55
40	4118 VERNON BLVD 11101	65-30 KISSENA BLVD	.	50
41	1146 44TH DR 11101	65-30 KISSENA BLVD	.	50
42	3544 STEINWAY ST 11101	65-30 KISSENA BLVD	.	48
43	2106 45TH AVE 11101	65-30 KISSENA BLVD	.	51
44	3421 37TH ST 11101	65-30 KISSENA BLVD	.	47
45	3114 38TH AVE 11101	65-30 KISSENA BLVD	.	53

Pros and Cons from using SAS and Google Maps

- Automated process
- No complex SAS coding
- Easy testing and checking results
- Usage for various commuting measures

- Substantial amount of time required for large inputs
- Process can't be left unattended for large inputs
- Some SAS code adjustments may be required to follow Google Maps updates
- Google Maps limitations

A Sample Commuting data

Commuting Time

Commuting Time

Table 1. Residency, Distance from College and Average Commute of Hunter Undergraduates (Fall 2008 & Fall 2009)

	All Students	Residency			Avg. Miles from Home Zip to Hunter			Avg. Commuting Time (in minutes) from Home to Hunter		
		New Freshmen	New Transfe	Continuing Students	New Freshm	New Transfers	Continuing Students	New Freshmen	New Transfers	Continuing Students
	27,982	3,818	2,933	21,231	34.1	46.4	16.4	79.2	88.2	55.9
	%	%	%	%						
NYC	**78.2**	**76.5**	**69.8**	**79.6**	**8.1**	**6.4**	**7.2**	**50.4**	**42.6**	**45.3**
- Bronx	8.1	7.1	8.0	8.3	8.1	7.5	7.8	48.2	43.9	46.1
- Brooklyn	25.4	25.6	21.6	25.8	9.4	8.0	8.8	55.3	49.9	53.1
- Manhattan	15.5	8.9	16.3	16.6	3.0	2.8	3.5	26.4	26.1	26.3
- Queens	25.8	29.8	21.8	25.7	7.2	6.3	6.7	48.8	43.5	45.3
- Staten Island	3.4	5.2	2.1	3.2	15.6	14.7	15.1	78.0	80.3	78.1
NYS	**9.7**	**12.5**	**12.8**	**8.8**	**34.5**	**42.5**	**34.6**	**103.0**	**110.3**	**102.2**
- Nassau	4.7	6.6	5.3	4.3	22.8	25.5	22.6	95.4	97.0	94.4
- Suffolk	0.9	1.2	1.1	0.9	51.3	50.5	46.5	133.8	128.3	125.7
- Westchester	2.1	2.4	2.7	1.9	22.9	26.5	21.0	72.5	76.4	70.0
- Other	2.0	2.2	3.7	1.8	71.6	76.1	71.5	157.8	164.5	161.3
USA	**3.0**	**3.8**	**6.5**	**2.3**	**614.6**	**526.7**	**280.1**	**832.7**	**752.3**	**322.1**
- New Jersey	0.9	1.0	1.3	0.8	22.1	20.6	14.5	102.4	88.1	74.5
- Other	2.1	2.8	5.2	1.5	825.4	659.1	417.9	1242.0	975.8	488.3
Foreign	**9.1**	**7.2**	**10.9**	**9.2**	**11.2**	**20.9**	**13.3**	**56.0**	**59.2**	**51.2**

Summarizing the Benefits of Using SAS and Google Maps for Commuting Times

- Effective tool for collection of commuting times and distance for the population of interest
- Allows to automate tasks for data extraction
- No need for special input formatting
- New version of SAS 9.2 contains new functions that allow user to compute geodesic distance

Thank you !