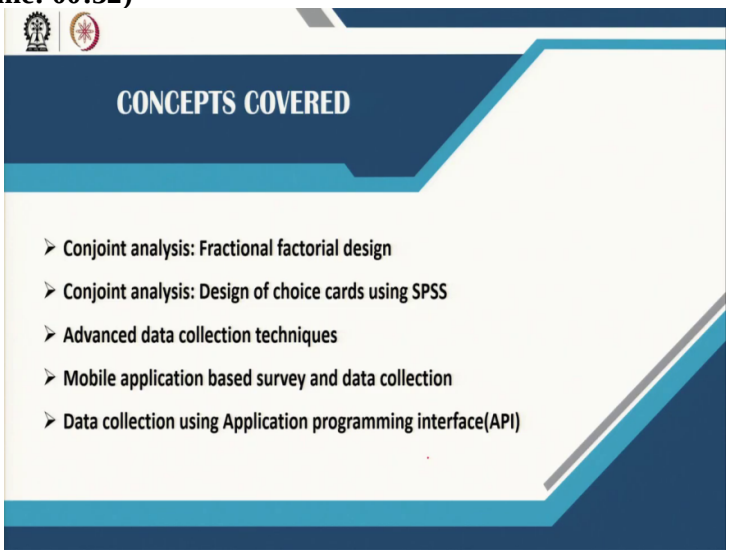
Urban Landuse and Transportation Planning
Prof. Debapratim Pandit
Department of Architecture and Regional Planning
Indian Institute of Technology - Kharagpur

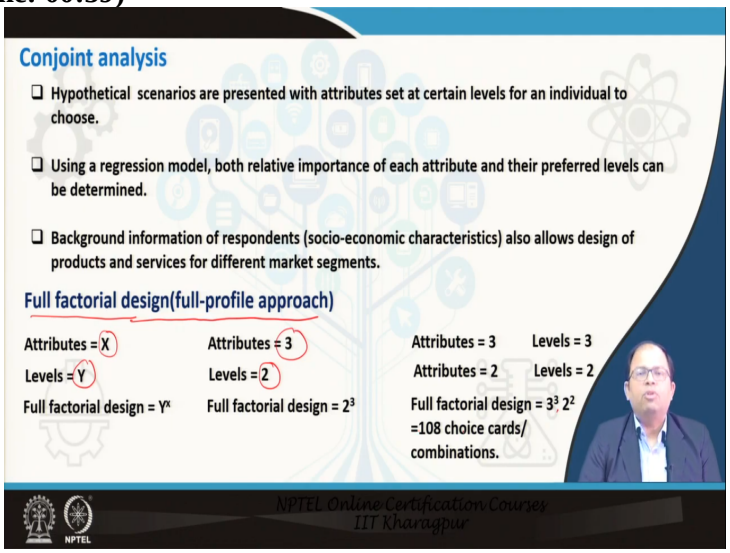
Lecture – 15
Transport Planning Surveys - Part 2

(Refer Slide Time: 00:32)



Welcome back. In lecture 15, we will cover the second part of transportation planning surveys. The different concepts that will be covered in this lecture are fractional factorial design which is done as a part of conjoint analysis. Then, design of choice set using SPSS software, advanced data collection techniques, mobile application based survey and data collection and finally, data collection using application programming interface.

(Refer Slide Time: 00:59)



Conjoint Analysis

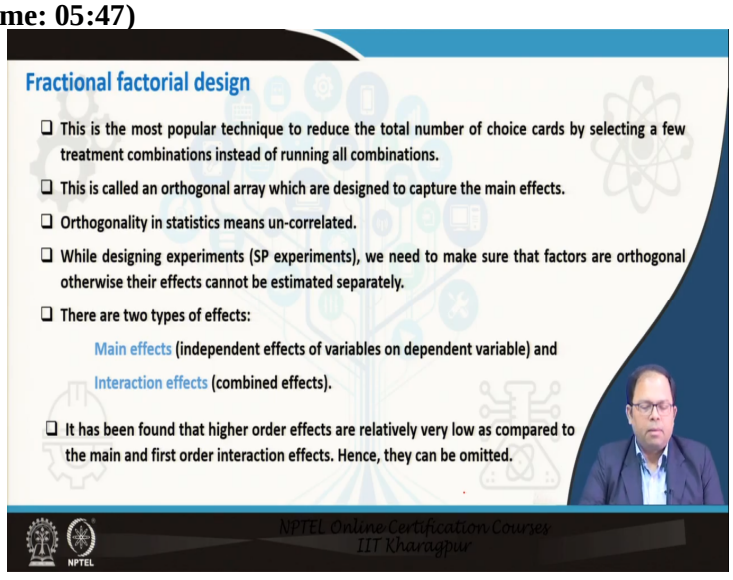
In the last lecture, we have seen that, in stated preference survey one of the methods is conjoint analysis, where hypothetical scenarios are presented to the actual person to choose as per his choice. And in each of these scenarios, a set of attributes is given with certain levels.

For example, for mode choice between bus and auto rickshaw, different attributes like bus fare, auto fare, travel time by bus, travel time by auto-rickshaw are given with varying levels. A combination of these attributes results in different scenarios, which are then chosen by the respondent. Then a regression model with this data gives the relative importance of each attribute and also their preferred levels. Along with this data, the socio-economic characteristics of the respondent can be used to design services for different market segments, as the influence of attributes vary across economic categories.

Full factorial design

There can be different attributes, like fare, travel time etc. as shown in the previous example. Each attribute can be set at different levels. For example fare can be Rs 5 or Rs 10. In a full factorial design or a full-profile approach, if there are X attributes and Y levels, the total number of combinations are Y^X. For example, with three attributes each having two levels, the number of combinations are 8. But for three attributes with three levels and two attributes with two levels the number of combinations would be $3^3 \cdot 2^3$, i.e. 108. For a surveyor to elicit response on so many combinations would be impossible. So there is a need to reduce the number of scenarios.

(Refer Slide Time: 05:47)

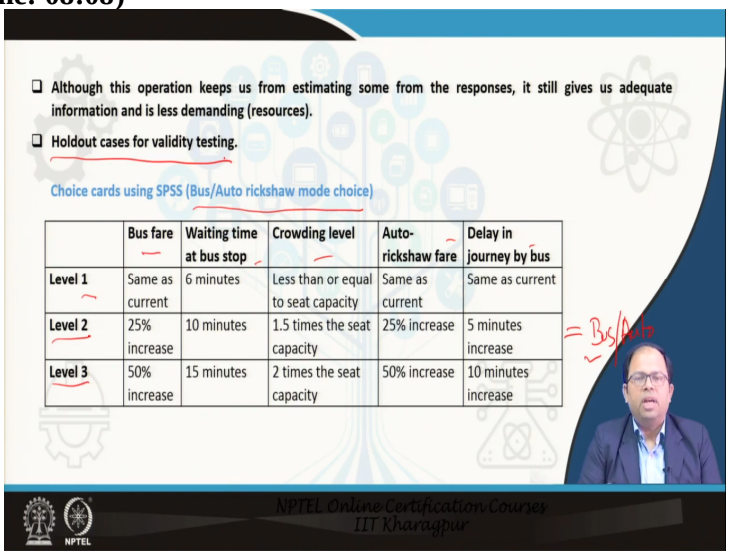


Fractional factorial design

One method for selecting a few treatment combinations instead of running all combinations is fractional factorial design. This is called an orthogonal array, which is designed to capture only the main effects. Main effect basically means whenever there are variables and there is a dependent variable, each variable has an effect on the dependent variable. So, independent effects of variables on the dependent variables are called the main effects. Whereas, when many variables affect the final choice of a dependent variable, it is called an interaction effect. Because of correlation, if two or three variables are playing a role on the dependent variable, in that case, it may be difficult to establish which one is having a main effect on that particular variable or what the actual contribution of that particular variable is. So, it is better to investigate where the cases are with only a single variable or the independent effect of a variable has to be only considered.

So, in orthogonal design an attempt is made to only capture the major effects or some interaction effects (the low level ones). The higher level interactions, where there are many variables working on that particular final dependent variable are avoided. So, orthogonality in statistics means uncorrelated, which is what needs to be achieve. Therefore, while designing stated preference experiments, one needs to make sure that factors are orthogonal otherwise their effects cannot be estimated separately. It has been found that, higher order effects are relatively very low as compared to the main and first order interaction effects and hence this higher order effects are not considered in a fractional factorial design.

(Refer Slide Time: 08:08)

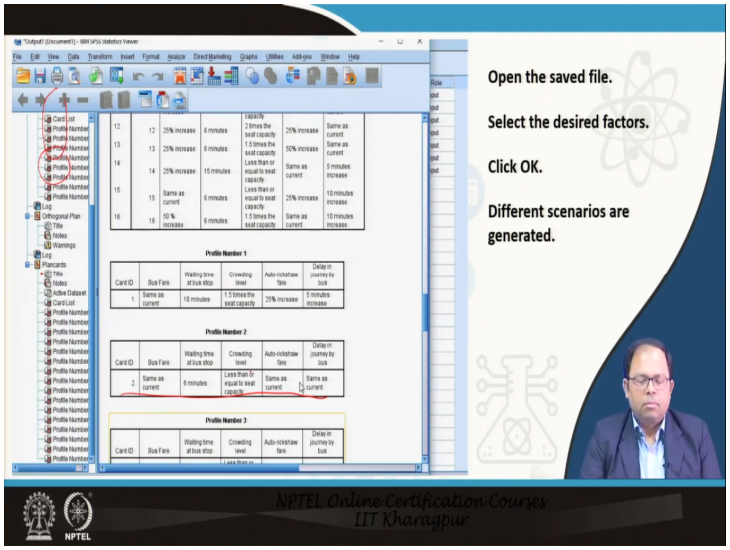


	Bus fare	Waiting time at bus stop	Crowding level	Auto-rickshaw fare	Delay in journey by bus
Level 1	Same as current	6 minutes	Less than or equal to seat capacity	Same as current	Same as current
Level 2	25% increase	10 minutes	1.5 times the seat capacity	25% increase	5 minutes increase
Level 3	50% increase	15 minutes	2 times the seat capacity	50% increase	10 minutes increase

For example, in a mode choice problem between bus and auto rickshaw with five attributes and three levels each, there can be many choice cards with a different combination of attribute levels. A fractional factorial design is used for reducing the number of such

combinations. IBM SPSS is one such software which can be used for this purpose. In addition to creation of this particular choice sets, holdout cases are also created. Holdout cases are a few more combinations than the ones that are finally chosen. Holdout cases are not surveyed but used for testing the validity of the design.

(Refer Slide Time: 11:13)



Following are the steps to create choice cards in SPSS software:

Step 1: Go to Orthogonal Design under Data tab`

Step 2: Press Generate

Step 3: In the dialog box fill in the variable names (should not contain spaces and should not start with numbers)

Step 4: Add all the other variables or attributes

Step 5: For each variable click on Define Values

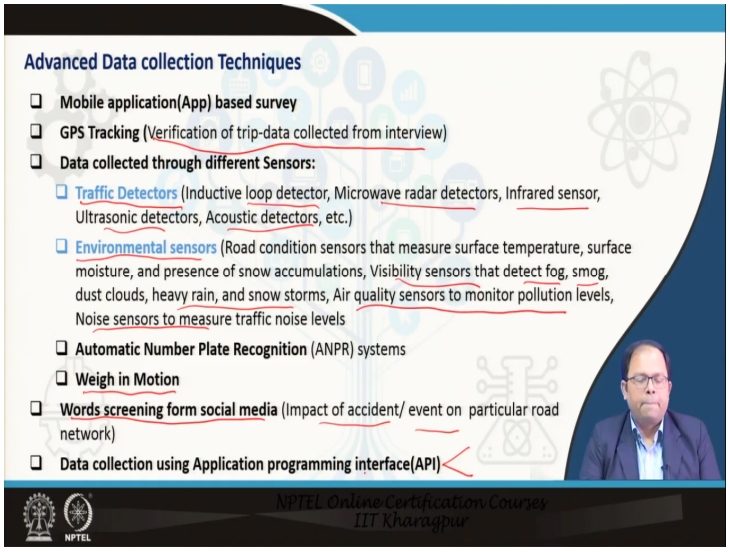
Step 6: Create the different levels for each attribute

Step 7: Save the particular file in a folder

Once the file is saved choice cards are automatically created.

In the example, instead of 3^5 combinations, there are only 16 combinations. Each respondent can be given 2 or 3 cards, and his/her choice recorded. Therefore, if 100 persons are given 3 cards each there will be around 300 samples on which a regression analysis can be done.

(Refer Slide Time: 15:21)



Advanced data collection techniques

These kinds of surveys are a completely different way of collecting transportation data. For example, GPS tracking using dedicated devices or smartphones. These kinds of data can be used for verifying trip data collected from interviews.

Data can also be collected from different sensors like traffic detectors (like inductive loop detector, microwave radar detectors, infrared sensors, ultrasonic detectors, acoustic detectors, etc) present in many roads for detecting the actual flow of vehicles, the speed of vehicles, and the number of vehicles passing through a particular point.

Similarly, there are environmental sensors which can tell us about road conditions, surface temperature, surface moisture or if snow accumulation has happened.

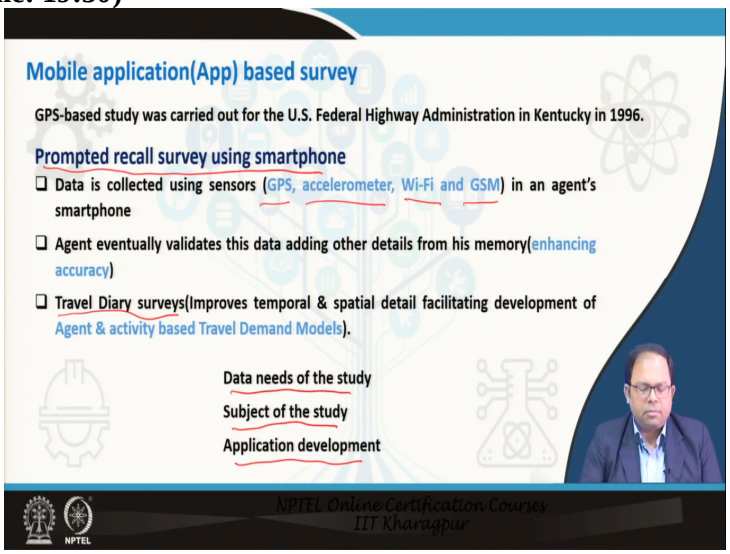
There are visibility sensors which can detect fog, smog, dust cloud, etc. Using this information, different warning message can be displayed to the users of a particular corridor. There are air quality sensors too, to monitor pollution levels and noise sensors to monitor sound levels.

There are automatic number plate recognition systems to identify vehicles breaking speed limits and signals. Weigh-in motion sensors can weigh freight vehicles climbing on to it.

In addition there is social media analytics, where social media feed can be crawled to detect accidents and congestions. Satisfaction with a particular transit system can also be gauged from the feed generated by social media users which can then be used to suggest improvements for the transit system.

Finally using application programming interface, one can directly access a database from another application. Access to such huge data is the key to big data analytics.

(Refer Slide time: 19:50)

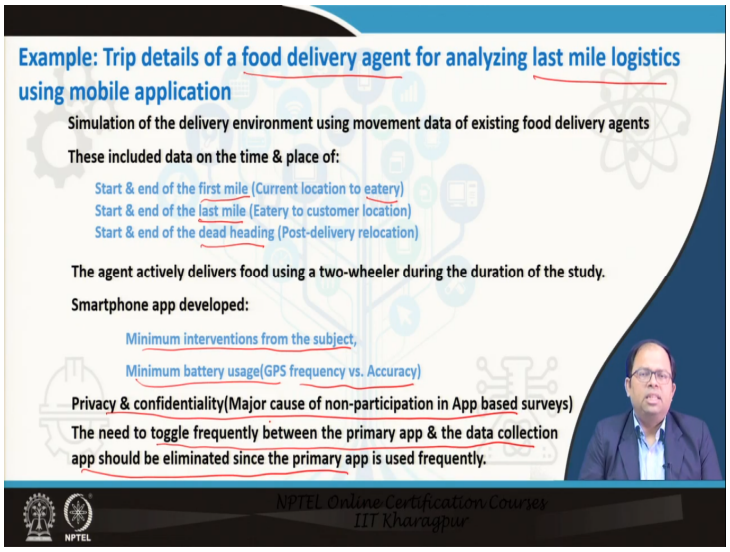


Mobile based GPS surveys

The first GPS-based study was carried out for US Federal Highway Administration in Kentucky in the year 1996. Prompted recall survey is a present day application of GPS based surveys. Locations visited by respondents are tracked using GPS enabled smartphones. A surveyor then verifies all these places and records the purpose of visit and other trip details in a face-to-face or telephonic interview. For this purpose the respondent has to install an application which collects information (like locations visited, routes traversed, speed) from the sensors (like GPS, accelerometer) and relays it to a remote database. With the proliferation of smartphones there is a huge potential of collecting such accurate data, provided privacy concerns are satisfied.

Thus, the three parts in the development of mobile based surveys are- assessing data needs, understanding the subject of the study, and the mobile application development.

(Refer Slide Time: 23:51)



Data needs of the study

In this example, an Android application is developed to remotely collect trip details of multiple food delivery agents for analyzing last mile logistics. Food delivery companies offer an online platform where customers can place orders for cooked food from nearby restaurants. In this case, last mile logistics deal with optimally assigning delivery agents to orders, thereby minimizing delivery time and cost. The objective of the study was to understand the pattern of last mile food delivery carried out by delivery agents- distances travelled before picking up an order, waiting times at the restaurant, distances travelled while delivering the order, deadheading (miles travelled after delivery), number of orders delivered per day, parking locations, etc.

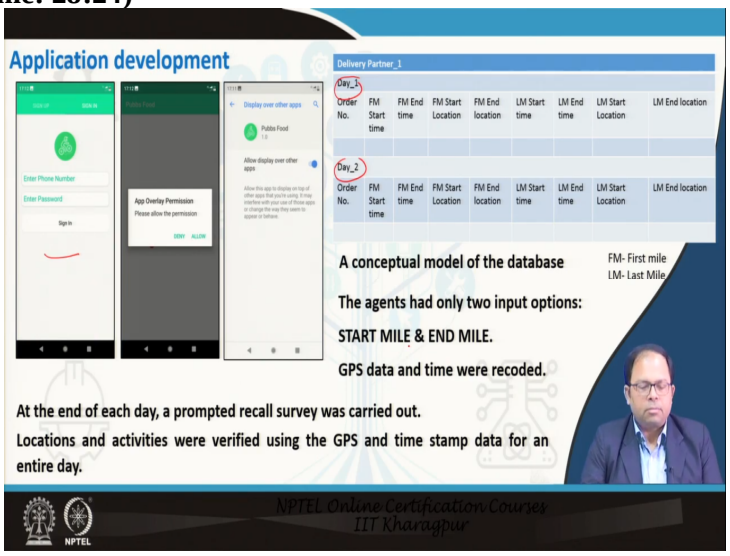
Subject of the study

The delivery agent uses a two wheeler to deliver orders. He is assigned orders via a smartphone application, which not only shows him his destination but also the shortest route leading to it. The data collection app needs to continually relay the delivery agent's location. As the agent is continuously dealing with the company's app(for route guidance) interventions are required for the data collection app to be minimal. Even then, the need to switch between the two apps can be difficult. Also, the agent needs to be tracked while only at work and not beyond delivery hours. Location tracking is an energy intensive process, draining a smartphone's battery very quickly; the additional energy requirement of the data collection app should therefore be very low.

Application development

Keeping the above requirements in mind in this example, an Android application was developed. The data collection process starts on opening up the app and ends when this app is closed. In the meanwhile, the application can access data from the GPS and accelerometer, to detect location and movement. To reduce energy requirements, the application does not continually relay the information to its remote database, but does so at regular intervals. Also, the frequency of collecting GPS location varies depending on whether the agent is static or has started moving. The application needed user intervention to identify restaurant and delivery location. This is achieved by a float button, which is overlaid on top of the primary app while the secondary data collection app runs in the background. Thus, need to toggle between the two apps was reduced. On reaching either a restaurant or delivery location, the agent marks the location by just accessing the floating button.

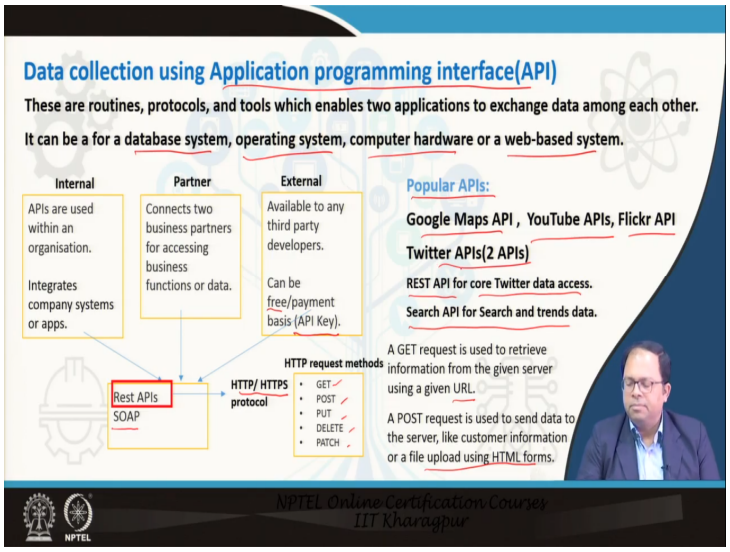
(Refer Slide Time: 28:24)



The information from multiple delivery agents were stored in a database. A conceptual model of the database is shown here. Since, the survey is carried out over multiple days, each agent's information is stored day wise. An ID is assigned to each order and the location and time at the start and end of the first mile and second mile is collected. First mile is the mile travelled before an order is picked up and last mile is the mile travelled before it is delivered.

Data collection via the app was followed by a prompted recall survey where the agent identified the restaurants from the location and the orders delivered were cross checked.

(Refer Slide Time: 30:11)



Data collection using Application programming interface (API)

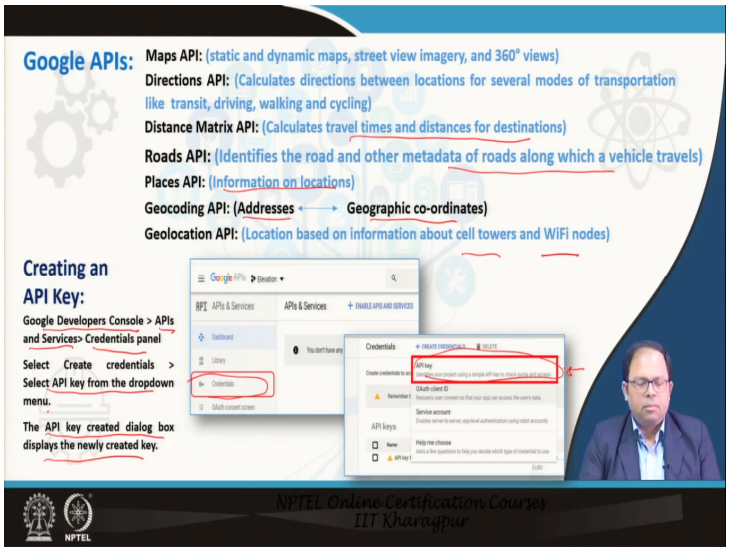
API refers to different routines, protocols and tools, which enables two applications to exchange data among each other. For example, the database of an application needs to be accessed via another application, third party applications need to access data from Google, etc. API interface can be for a database system, operating system, computer hardware or web-based system. Since the applications are different, a common interface is required between them.

Google Maps, YouTube, Flickr, Twitter give access to their data via an API key. Some of these are payment based. Along with APIs, there needs to be a set of rules which determine how different systems talk to each other. Rest API, or restful API is a system architecture which says how different data can be exchanged between two programs. SOAP architecture is one such architecture. And HTTP and HTTPS are the two protocols.

There are different HTTP request methods such as GET, POST, PUT, DELETE, and PATCH. A GET request is used to retrieve information from the given server using a given URL and POST request is used to send data to the server like customer information, or a file upload using HTML forms.

Twitter has got two APIs. One is the REST API for core Twitter data access using which one can analyze Twitter feeds (what people are tweeting about). And the other is Search API for accessing search and trend data..

(Refer Slide Time: 35:06)



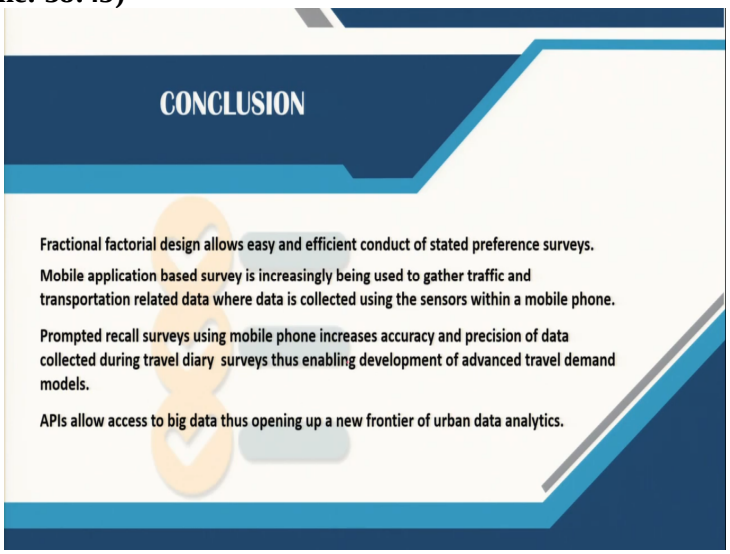
Some of the popular Google APIs, like Maps API, provides static and dynamic maps, street view imagery, and 360 degree view. So if a program requires data about certain street view imageries, certain 360 degree views of certain areas, one can access the maps API. Direction API shows the way a particular transportation mode can travel between 2 locations. For example, one can determine the starting and ending point of a trip or the actual route a person follows during a particular trip using this API. These systems can be integrated with existing land use transportation models to achieve better visualizations as well as to get feedbacks like existing travel time or travel conditions in a particular corridor. Google API like distance metric API calculates the travel times and distances between an origin and destination. This can become a determinant in a mode choice model or trip assignment or a route choice model. Roads API, gives access to metadata of roads. Place API gives information on locations. Geocoding API can give access to geographic coordinates of addresses or vice versa. Geo-location API gives real time location of a particular person using cell towers and Wi-Fi nodes, which can be an alternate to GPS.

A Google API key can be created from the Google Developers Console as is shown in the slide. Payment is required beyond a certain number of requests.

(Refer Slide Time: 38:26)

These are the references that can be followed. A free version of IBM SPSS is also available.

(Refer Slide Time: 38:45)



Fractional factorial design allows easy and efficient conduct of stated preference surveys. Mobile application based survey is increasingly being used to gather traffic and transportation related data where data is collected using the sensors within a mobile phone. Prompted recall survey using mobile phone increases accuracy and precision of data collected during travel diary survey, thus enabling development of advanced travel demand models. And finally, APIs allow access to big data that is opening up a new frontier of urban data analytics. Thank you.