**Data Analytics Task 1**

**By *Pranjal Rastogi***

**Question 1: What is Data Analytics?**

**Answer: -** *Data analytics* refers to qualitative and quantitative techniques and processes used to enhance productivity and business gain. Data is extracted and categorized to identify and analyze behavioral data and patterns, and techniques vary according to organizational requirements.

It is the process of examining data sets in order to draw conclusions about the information they contain, increasingly with the aid of specialized systems and software. Data analytics technologies and techniques are widely used in commercial industries to enable organizations to make more-informed business decisions and by scientists and researchers to verify or disprove scientific models, theories and hypotheses.

Data analytics is primarily conducted in business-to-consumer (B2C) applications. Global organizations collect and analyze data associated with customers, business processes, market economics or practical experience. Data is categorized, stored and analyzed to study purchasing trends and patterns.

**Question 2: - What is the difference between data science and data analyst?**

**Answer: -** The *Data Analyst* typically runs queries against new data to find trends important for the organization and to help prepare data for the Data Scientists. Data Analysts are typically very good at SQL as well as being knowledgeable of the core metrics an organization deems important. They can also write scripts and produce intuitive visuals.

Whereas, the *Data Scientist* is primarily tasked with building models using machine learning. These models are expected to engender an organization’s software with product features that predict and explain; making applications adaptive.

The quality of a Data Scientist’s models depends directly on how well they understand and prepare data; thus, they will work with the Data Analyst when it comes to understanding and preparing data to build better models.

**Question3: - What is need of Data Analytics in governance?**

**Answer: -**

Certain government agencies face a big challenge: -

Tighten the budget without compromising quality or productivity. This is particularly troublesome with law enforcement agencies, which are struggling to keep crime rates down with relatively scarce resources. And that’s why many agencies use big data analytics; the technology streamlines operations while giving the agency a more holistic view of criminal activity.

Although decision makers in businesses are limited, they are a diverse set in government. The government has tremendous data in legacy databases and forms that must be curated and migrated for new-age analytic tools. Data sharing between departments and across ministries is a challenge, provided the existing jurisdictional boundaries. Huge data and associated analytics are advantageous in numerous areas

For a country like India that owns the largest complex of data gathered via digitalization of records for purposes like IDs, passports and payment of subsidies Big data is characterized by its volume, variety and speed and the analytics involves its processing in a cost-effective way in order to draw conclusions for their useful application.  
There are a number of areas where huge projects have been implemented, like Aadhaar, passports and the inception of MCA 21.

Besides it can be used to solve problems including traffic problems in cities, informed policy making, enabling security to people and society.

Big data analytics also merges into fields like deep learning, machine learning and artificial intelligence (AI) and thus, has tremendous possibilities  
With IoT (Internet of Things) now coming onto their own after AIT and ITeS (information technology enabled services), a whole new world is opening up for data with things like sensors which can be worked up to reduce problems of traffic, increased monitoring security etc.

**Question4: - Define a use-case according to your own understanding where Data Analytics can be used in Governance?**

**Answer: -**

***Use Case : Video Analytics for Surveillance and Safety***

**Introduction:** Security is one of the primary concern for any city and takes a big chunk of budget especially for big cities. Though Law enforcements are doing their best to keep crime from happening but with rising technology it is preferred for governments to take next step towards citizen safety.

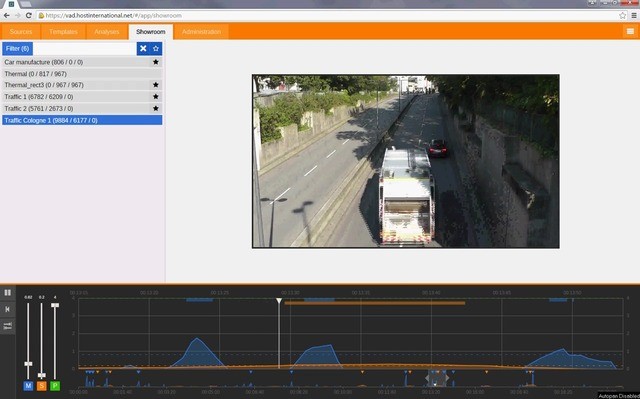
So many places are under CCTV surveillance now a days and whole lot of data is available, thus data analytics can be applied to form some predictive algorithm in order to find traits of a robber, burglar etc. so as to trip off alarm and identify suspects.

**Brief:** Protecting infrastructure goes beyond predictive maintenance, and frequently people need protection *from* infrastructure. This was first implemented in protecting oilfield infrastructure through deployments of cameras alongside motion and radar sensors.

In surveillance contexts, the detection of anomalies is crucial, and can be done through a novel mechanism: “We let the machine try to figure out what the state of the feed is and derive normal or abnormal things over time. We create anomaly indices, and if something pushes past threshold it’s counted as an anomaly.”

Human operators monitoring the sensor feeds can even adjust the threshold for what counts as an event worthy of their attention, as in the following example of a traffic management system based on video analytics.

An example picture showing similar data analytics implemented by AGT.



*Example: Adjusting anomaly indices in AGT’s video-based traffic management system*

The line in the graph at the bottom represents the threshold of significance set by the operator, while peaks indicate anomalous events (a stalled car, a car moving in the wrong direction etc.) that pass the threshold and hence require the attention of a human being.

Because this technology is based on machine learning, users don’t have to define complex rules to determine what counts as an event.

Instead, the software itself figures this out: “If you use this in traffic, the analytics will learn that there are four lanes, two in each direction, and typically the car flow goes on the left side from the top to the bottom and on the right side from the bottom to the top. If the analytics detect a strong movement from left to right in the video signal, that’s an indication of an accident and the analytics will trigger an alarm.”

Such technologies are useful for adjusting operations, as well as for ensuring safety. AGT’s video analytics are also being used to manage crowd movements in queues and crowd density at events.

**Steps Toward Implementation:**

The use cases we’ve examined requires following steps:-

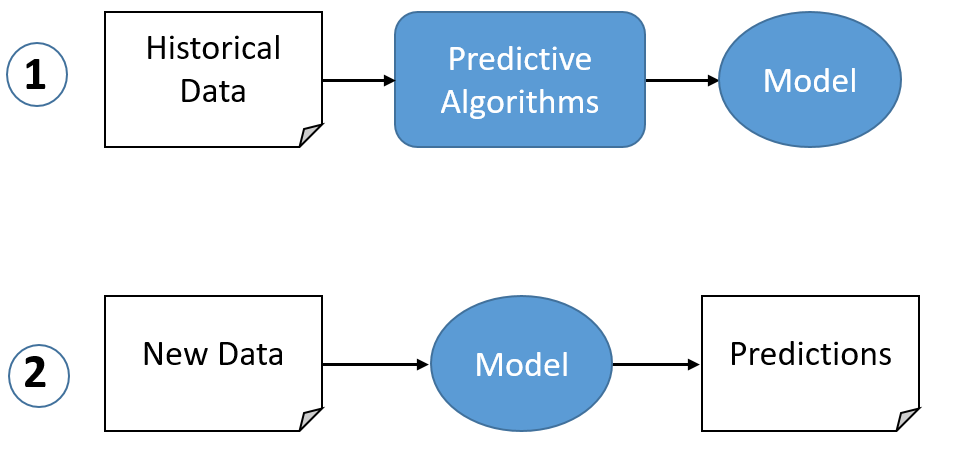
**1)Data Provided and Data Reconciliation**

The data can be provided by the Police Service in two lots:

• An ACCESS database of incidents in the given time period

• An EXCEL spreadsheet of locations of surveillance cameras

The EXCEL spreadsheet provided the map location of each camera, given in coordinates from GPS measurements, in degrees and minutes of angle, with the minutes specified to three decimal places.



**2)Data Selection**

The first step in the analysis is to define ‘target areas’ and select all the incidents in those areas before, during and after the implementation of the CCTV camera system. This can then be cleared and selected using a language and basic data can be retrieved through some algorithm such as *regression.* This can be further improved by letting it run for long time.

**3)Summary Statistics**

For each target area considered, summary statistics can be calculated - the total and the average monthly number of recorded incidents and associated standard deviation - for the selected areas during the implementation period and during the selected period before implementation. Simple percentage changes in the average monthly number of incidents before/after implementation were calculated, with positive values indicating a decrease in crime after implementation (desired effect) and negative values indicating the opposite effect (a value of zero indicating unchanged level of crime).

**Strategic takeaway:** Expensive industrial infrastructure can be managed and protected through IoT analytics. This includes video analytics as well as sensors. Moreover, new technologies supporting the connected city will also be largely enabled through IoT analytics.