**Introduction to the Problem Statement**

Say that you are working as an analyst at Smart Consultancy. Your company offers analytical solutions to various clients across the various domain. One such client of your company is the state's open-pit mine. Their product head approaches your company with a problem faced in the mine. The mine has operations such as the digging of ore and crushing of ore into a finer composition where this ore gets transported between the diggers and crushers using transportation trucks.

 The open-pit mine is facing problems of inefficient production and is losing customers' trust as they are not able to meet their demands even though there has been no surge in demand. The mine operators are sure that this is a problem at their end and want to first understand their tracking and operations on the field in detail, as it is not possible to keep track of all operations at one go due to the vast expanse of the mine.  
As an analyst, you have been approached to build a smart live monitoring system. You, as an individual, have been assigned the task to understand the key metrics explained by the client and apply your knowledge of supply chains to get track inventory and quantity available and understand what is happening on the ground at each location. You have been given some key metrics to capture and what operations are going on in the mine. Your job as an analyst is to perform the following 4 tasks:

* Understand the problem faced by the mine and take an understanding of the data shared by the client.
* Since the client is not very tech-savvy, you need to prepare a cleaned dataset using the dataset provided by him.
* Analyse the data using the MySQL Workbench 8.0 and prepare a live tracking system using some set of dashboards using either Tableau or Power BI.
* At last, present the findings or key insights to your senior technical manager using a ppt of not more than 8 - 12 slides who wishes to understand the key workings of your analysis along with the insights derived out it.

*You will get more details on the data, processes on the mine and the key metrics to be calculated in the upcoming segments.*

Till then, let’s listen to Gautam as he introduces the problem statement in the upcoming video.

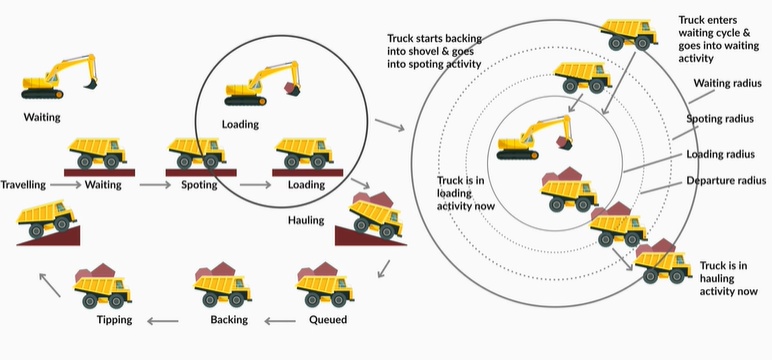
In this session, we will cover the following points, which will help you with designing the model:

1. Understanding the digger–crusher operation
2. Understanding the data provided by the customer
3. Looking at the steps needed to create the dashboards
4. Learning about the key metrics that needed to be presented

# Digger–Crusher Cycle

Before we dig deeper into the data for the problem, let’s watch the upcoming video to get a deeper understanding of the truck, digger and crusher operation in open-pit mining from Gautam. This will help you understand some key on-field operations on the open-pit mine.

So, as Gautam explained in the video after drilling and blasting are completed, diggers excavate the mined ores, where they start the crowding operations, i.e., gathering the ore into stockpiles. This ore is then are loaded into dump trucks, which take these ores back to the crusher sites where they are crushed into granular pieces. The empty trucks go back to the digger sites to collect additional ore.  
If empty trucks are not queued at the digger sites, then the digger operation will have to stop until the next empty truck arrives. Similarly, if there is no loaded truck at the crusher site, the crusher operation is stopped until a loaded truck arrives and dumps additional ore. Thus, in a typical scenario, it is essential to ensure that sufficient trucks are assigned at diggers and crushers so that their operations are running smoothly. This will ensure that continuous production is maintained and is maintained at a certain rate.



Truck Cycle

In a typical Truck Cycle, an empty truck lined up at a digger (Waiting Stage) is first parked at the Loading position (Spotting), where it gets loaded by the digger. After the truck is loaded at the digger, it drives away from the digger and moves to the crusher. This process is called Hauling. Here, it is called hauling because it carries a load to the crusher. At the crusher, the truck is again queued, waiting for its turn.

 As soon as its turn comes, it backs into the crusher which is called Backing and at this position, it would offload the ore into the crusher and this is called Tipping. After this, the truck becomes empty and heads back to wait for loading at a digger (Travelling). This cycle from Waiting, Spotting, Loading, Hauling, Backing, Tipping, Travelling and back to Waiting is called the Truck Cycle.

 Similarly, diggers also operate in a cyclic manner. Each digger has a swing arm, which carries the load in a clockwise direction in order to gather the ore and add to the stockpile. This is called Crowding. Next, the digger would place the bucket load, which is called Load, into a truck. This requires multiple cycles to load the truck while using buckets to load.

 Now that you have understood the digger and crusher cycles, let's look at the data provided in the next segment.

**Data Explanation**

In the problem statement, you will be provided with 3 primary datasets.

* Cycle data consisting of the movement information of the equipment, duration of the cycles and much more
* Location data consisting of the location information of the equipment and the other similar attributes.
* Delay data consisting of the delay in service time of the on-ground equipment for both planned and unplanned delay schedules and other important attributes related to delay.

Let's start understanding the dataset in detail one by one.

 In the upcoming video, we will look at the cycle data description that has been provided by the client.

PauseYou can download the dataset below. We request you to go through the above video carefully. This is the entire cycle data received from the client. You will have to prepare a proper dataset based on the information needed to prepare the dashboard.

**Cycle Data**

**Download**

So, now that we have seen the cycle data, in the next video, we will look at the location data, which shows the location of each crusher and digger.

ReplayYou can download the location data from below.

**Location Data**

**Download**

Finally, in the forthcoming video, we will look at the delay information provided by the customer.

PauseYou can download the delay data from below.

**Delay Data**

**Download**

In addition to the above data, there is an additional dataset that you will have to derive. And that can be derived from the cycle data, which is essentially the movement data of the equipment on the open-pit mine. It should have the following details.

* Machine Name
* Source Location
* Destination Location
* Payload
* Start Time
* End Time
* Movement Id

 Here, is the data dictionary for reference for all the above 3 datasets. Please go through the dictionary carefully to understand the different variables in the dataset.

**Data Dictionary**

**Download**

Now that you have understood the data, it is your job to create the cleaned data for the key metrics that you are required to calculate. You need to download the above CSV files and load them in MySQL Workbench.

 After having done that, you will have to create stored procedures for the following

* Equipment Master
* Equipment Type Master
* Location Master
* Location Type Master
* Stored Procedure for Cycle Data
* Stored Procedure for Delay Data
* Stored Procedure for Movement Data
* Stored Procedure for OEE (you will understand what is OEE in the upcoming segment)  
   After having completed the stored procedures above, you will then perform analysis on the data to create suitable visualisation dashboards in Tableau or Power BI.  
    
  Let's understand the next steps in detail in the upcoming segment.

**Data Processing**

We will begin this segment with a video to understand how we can use the data provided in the previous segment to build an analytic solution.

PauseAs you have studied in exploratory data analysis, the first and the foremost task is to clean the data by identifying relevant columns, removing outliers and filling in the missing values for those columns. Then you need to prepare a master table, which has all the key columns. As discussed in the previous segment, there will be the following master tables:

* Equipment Master
* Equipment Type Master
* Location Master
* Location Type Master

After this, you need to prepare process tables, which are formed with interrelations between multiple master entities (for the process tables, you may need to perform outlier detection and fill in the missing values again). After this, you need to calculate the key metrics from the process tables, which will give you more insight, and by presenting the data in a visualisation dashboard and adding additional sequential patterns (SPs), that show the on-ground performance.

 Stored Procedures for

* Cycle Data
* Movement Data
* Delay Data
* OEE Calculations

Let's now understand the key metrics that you have to calculate in the upcoming segment.

**Key Metrics**

In the upcoming video, we will look at the key metrics that you will need to prepare the live tracking dashboard.412

So, as Gautam explained in the video, these dashboards will be available to both high-level executives and on-ground technical staff. Therefore, key metrics should include:

1. Equipment Availability
   1. Number of equipment operating in the field
   2. Number of equipment under maintenance
2. Production Rates
   1. Production versus plan
   2. Top-performing equipment
   3. Poorly performing equipment
3. Efficiency of Operations
   1. Overall equipment effectiveness (OEE): Availability, performance, quality

To calculate OEE, you need to first calculate the availability, performance and quality metrics as shown below:

Availability    = (Net Available Time – Down Time)Net Available Time×100  
  
Performance = (Operating Time – Speed Losses)Operating Time×100

Quality             = (Net Operating Time – Defect Losses)Net Operating Loss×100

Finally, OEE is given as the product of the above three metrics, as shown below:

 OEE=Availability×Performance×Quality

 On the basis of the dataset that you prepare, capture these metrics using the live tracking systems and build dashboards using Power BI or Tableau.   
  
You will need to build 5 dashboards.

1. Executive Dashboard: Level 1 Dashboard with all the details in one.
2. Equipment Level Production
3. Efficiency, Accuracy and Quality of the Equipments
4. Key Metrics of Lowest Performing Equipments
5. Top Performing Equipments

**Note:** The dashboards above should have the right visualisation chart depicting the information drawn from it. Recall the concepts taught in the visualisation modules and also recollect the right charts to be used in the correct scenario based on the data storytelling module.

Finally, towards the end of the project, you will be building a PPT of not more than 8 to 12 slides describing your work to a senior technical stakeholder. You would want to explain to him the details of your analysis plus the insights derived out of it and give some recommendations to the client for the solution that he is looking for with regard to his open-pit mine problem.  
  
It is your job to figure out what details to put in the PPT based on the information provided above and also considering that it is for a senior technical stakeholder. You will also have to recall the elements of data storytelling and incorporate those in your PPT. Also, make sure that the PPT is consists of the best practices taught in that module.  
  
In the next segment, you will understand the evaluation rubric for this project.