**Construction of OLAP Cube through Power Pivot in Excel**

**Group Members:**

* **MUHAMMAD SUMAIR – 22995**
* **KHANSA JUNAID – 23074**

**About the dataset:**

**Chosen Dataset**: New York City Buss Breakdown and Delays

The Bus Breakdown and Delay system tracks real-time information from school bus company in the field and records incidents of delays. This information is used to notify OPT and provide updates to inquiring parents also the data is recorded by the school bus company personnel.

**Problem #1: Understanding the reasons of Bus delays.**

* **Facts:** Bus breakdown ID, how long delayed.
* **Dimensions:** Route Number, Breakdown or running late, Occurred on.
* **Useless Variables:** School year, run type, Bus No, Schools serviced, Created on, Boro, Bus company name, Number of students on the bus, Has contractor notified schools, Has contractor notified parents, Have you alerted OPT, Informed on, Incident number, Last updated on, School age or prek.

**About this problem:** In this problem statement we will analyze the reasons that have the greatest number of delays. We will also consider the time of delay, time of occurrence, route numbers and the number of buses that are involved in the problem. Keeping in mind all these factors, we will analyze the most frequent issue.

**Process diagram:**

**Group the data by time to see how delays change over time, such as during peak hours or weekends.**

**Drill down further to see what types of delays are causing the most issues, such as traffic etc.**

**Start with Total number of bus breakdowns and delay time for each route.**

**Finally, rank the bus routes based on their overall delay performance, with a focus on delay time and number of breakdowns.**

**Charts:**

Fig 1

Fig 2

How long delayed

Fig 3

Fig 4

***For a better display we have applied filter on the route number and the time it occurred etc.***

**Recommendation:**

If we understand the dataset provided to us, we can construct the graphs or charts as shown above. We can read these charts easily by looking at both the axis and the color schemes that are the key.

If we look carefully at figure 1, we can see that almost every bus faces a delay of around 30-40 mins. Furthermore, most of the delays happen in the morning, which is either school time or office time. If we go into more detail, we can see a pattern in the charts, that is, most of the delays are due to buses running late. Figure 3 shows that there are a lot of instances where buses who are running late face the most delays as compared to the bus facing the breakdowns. Similarly, figure 4 two sections, breakdown and running late, where breakdowns are very less on different routes that cause any problem while on the other hand bus running late face most of the issues such as heavy traffic, late return from the field trip etc.

To solve the problem, there are some suggestions that will help everyone, which is that buses should leave on time or sometimes early, so that the traffic is avoided. Secondly, the bus companies should maintain their vehicles so that their buses are not delayed. Although breakdowns have caused a little delay when compared to buses running late, still it would be highly efficient if there were no minor breakdowns.

Moreover, transportation authorities should make sure that heavy traffic is not allowed during the school hours like in the morning or in the afternoon when the school gets over. Furthermore, the bus companies should give knowledge to their drivers about the alternate routes which are comparatively less packed than the usual through gps systems so that the delays are minimized.

**Problem #2: Understanding the company’s performance.**

* **Facts:** Bus breakdown ID.
* **Dimensions:** Breakdown or Running late, Bus Company Name.
* **Useless Variables:** School year, run type, Bus No, School Serviced, Created On, Boro, Number of students on the bus, Has Contractor notified schools, Has Contractor notified parents, have you alerted OPT, How long Delayed, Incident number, Last updated on, School Age or Prek, Occurred On, Informed On, Route Number, Reason.

**About the Problem:** This dataset can also be used to analyze the performance of different Bus companies. The transportation authorities can evaluate the level of service offered by each company and decide on future contracts by looking at the data on the number and types of occurrences recorded by each company.

**Process Diagram:**

**Select “Bus Company Name” dimension and then select the desired company.**

**Select “Breakdown Id” and “Breakdown or Running late” dimensions.**

**Count the number of events for each company.**

**Explore the trends in the data and also know about the service provided by each company, this will ultimately help you make suitable decisions for your future contracts**.

**Use the resulting data to analyze the frequency and type of events reported by each company**.

**Charts:**

Fig 1

**Recommendation:**

From the data set given we have constructed the graph shown above based on our analysis of understanding Bus company performance. From the graph, the events(Breakdown or Running late) occurred due to several reasons can be easily counted. As seen Bus company with the greatest number of delays is Leesel Transp Corp, with the reason being heavy traffic. Although other causes are also part of the bus being delayed like mechanical problems, flat tires, won’t start etc, but the most prominent reason for this specific company is heavy traffic. Buses that belong to other companies also get delayed due to numerous reasons but buses that belong to leesel corp are mostly likely to be delayed and due to traffic reasons. This is the cause of concern as why other busses are not subjected to traffic as much as them.

School authorities should look into this matter and should raise this concern to the vendor regarding their service as the bus being late badly affects students’ attendance and time. To make it better, companies can give knowledge to their drivers about alternate routes and also train them to efficiently check maps to avoid areas full of traffic. Also, as breakdowns are also one of the reasons for bus delays, companies should do inspections and should also maintain their vehicles timely to minimize mechanical problems. Though this constitutes a very small percentage for the bus being late but if these small percentage is also minimized it can have a very large impact in lessening the Bus delaying time.

Based on the analysis, school staff can make some suitable decisions about who to contact for future contracts and providing the resources to the company that are providing the highest level of service.

**Problem #3: Understanding the accuracy of event reporting.**

* **Facts**: Route Number
* **Dimensions**: Occurred On, Informed On, Bus Company Name, Has Contractor notified parents.
* **Useless Variables**: School year, run type, Bus No, School Serviced, Created On, Boro, Number of students on the bus, Has Contractor notified schools, Have you alerted OPT, Breakdown ID, How long Delayed, Incident number, Last updated on, School Age or Prek, Breakdown or Running Late

**About the Problem**: This dataset can also address the analysis of the accuracy of events reported by bus staff. The transportation authorities can evaluate the accuracy of the event reporting system and spot areas for development by looking at the data on the quantity and time of incidents recorded.

**Process Diagram:**

**Select the dimension “Has the company notified parents”, either yes or no.**

**Select the dimension “Route number” and the desired route.**

**Select the dimensions “Occurred On” and “Informed On”.**

**Select the dimension “Bus Company Name ” and the desired company.**

**Used to identify areas of improvement, such as reporting incidents and addressing specific routes.**

**Use these results and, if you find a No then the company is not fully accurate in reporting.**

**Use this data to identify trends also can be used to assess the accuracy of the events reporting based on different routes and different bus companies**.

**Charts:**

Fig 1

**Recommendation:**

From the graph shown above we can analyze the accuracy of event reporting. As we can see graph is specifically about the whether the contractor has notified the parents about the occurrence of incident or not. As seen Pioneer named bus company has frequently not notified the parents about the status of the bus. Though other vendors has also not frequently not notified the parents about the delaying but Pioneer has the highest probability. Also there are many bus vendors who have responsibly played their role but schools associated with Pioneer and other bus vendors who has not informed the parents, must have complains from the parents about this problem.

To solve this, company must make sure that quality ensured measures are taken so that events are accurately and promptly reported. To make this happen they must train their staff to inform about their status to the school and parents, also immediately notify parents if they are running late due to whatever reasons. Also transportation authorities can provide a more robust system for reporting the event on time. Also, company should also recommend regular audits of event reports so that any type of event whether it is delaying, or breakdown must be informed promptly to both school and parents.