Excel Challenge

1.

Conclusion1.

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

According to the pivot chat, we can see the category “theater” has the most of data and the category “music” has the second highest number of occurrences. However, the percentage of successful for “theater” is less than “music”, which means the chance of successful for “music” is larger than “theater”.

Second, the percentage of failed for “theater” is more than “music”, which means the chance of failed for “music” is less than “theater”.

Finally, the number of live are the same between those two categories, but the percentage of live for “music” is larger than “theater” due to the difference of the total.

So, in my opinion, the campaign made on “music” is more efficient than made on “theater”.

Conclusion 2.

一張含有 坐, 彩色, 填滿, 光 的圖片

自動產生的描述

一張含有 螢幕擷取畫面, 標誌 的圖片

自動產生的描述

Due to the most of campaign made on the parent category “theater”, the sub-category ”plays” would be the most data as well. However, we can see the number of failed is largest even the number of successful is larger than the total of others.

Ee can see the parent-category “music” is more efficient than “theater”, and for the sub-categories in “music”, “rock” has the largest number of successful and never get failed.

In my opinion, the campaign made on rock music has more chance to be successful than others.

Conclusion 3.

一張含有 文字, 地圖 的圖片

自動產生的描述

According to the trend line chat, we can see the number of successful on May is the most, and the number on March is the second peak before the number drop, and the number on December is the lowest, even less than the number of failed.

The number of failed, from Feb. to Apr. are lower than other months, and the number of canceled seems not different between the months but the change between Feb. to Jun would not be variable.

In my opinion, the campaign made in Feb. would be more efficient than other month because the number of successful is high, the number of failed is low and the variance of canceled would be low too.

2. limitations of this dataset

The numerical variables which can be analyzing or determining as the factors in this dataset is only 3, and that is hard to know how they define the “state” as successful, failed. We can only know that ”what the story is” by the summary of data but not the reasons “what make the story”.

3. some other possible tables and/or graphs that we could create

Maybe we can put “country” as the filter or another variable in the pivot table by month and see what the result is in every country, and we would be able to know how to make the strategy in the specific area.

Bonus Statistical Analysis

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

一張含有 螢幕擷取畫面 的圖片

自動產生的描述

I draw the Median as the Red line and the Mean as the Black line. The most of data is close to the Median, so the Median for those two variables would be more meaningful as the data center. And for variability, successful has larger range than failed and the normal distribution chart shows successful has more variability than failed.

It makes sense for failed, because whenever the number of backers increase, the chance of failed will drop but there might be some other reasons make the campaign failed.

And for the successful, it is hard to say make sense, because even though the number of backers increase, the chance of successful drop after meeting the mode. That is not the same result as the failed, it should be other factors which effect the chance of successful.