

DESIGN THINKING

1. Bussiness Understanding

Dalam menyelesaikan permasalahan ini, perlu diketahui daerah-daerah mana saja yang memiliki penumpang paling banyak berdasarkan data-data yang tersedia. Setelah mengetahui daerah mana saja yang memiliki penumpang yang banyak, maka dapat diketahui rute-rute yang membutuhkan penambahan armada.

Sedangkan di beberapa lokasi terdapat penumpang yang sedikit sehingga terjadi ketimpangan antara armada dan penumpangnya. Maka dari itu, perlu diketahui daerah mana saja dan faktor apa saja yang mempengaruhi sedikit atau banyaknya para penumpang bus sekolah di setiap area. Dengan mengetahui berbagai macam faktor, maka dapat digunakan untuk menentukan strategi sosialisasi atau penambahan armada.

2. Data Understanding

Untuk bisa menyelesaikan permasalahan ini diperlukan beberapa data yaitu :

1. Data Penumpang Bus Sekolah Tahun 2017

Data diambil dari

<http://data.jakarta.go.id/dataset/data-penumpang-bus-sekolah-kpi> yang terdiri dari 12 data penumpang bus sekolah setiap bulan di tahun 2017. Berikut ini adalah variabel-variabel yang terdiri dari :

1. type_operasi
2. daerah_operasi
3. area_operasi
4. jumlah_bus
5. jumlah_penumpang

Berdasarkan variabel tersebut, terdapat tiga data object atau kategorikal yaitu type_operasi, daerah_operasi dan area_operasi. Sedangkan untuk tipe data integer terdapat dua variabel jumlah_bus dan jumlah_penumpang.

Berdasarkan data tersebut bisa dibandingkan area mana yang perlu penambahan armada karena tidak mencukupi banyaknya penumpang yang ada.

2. Data Jumlah Armada Bus Sekolah 2017

Data diambil dari

<http://data.jakarta.go.id/dataset/data-jumlah-armada-bus-sekolah/>.

Berdasarkan data tersebut terdapat beberapa variabel yang dapat digunakan, yaitu :

1. Jenis Armada
2. Merk
3. Tahun Pengadaan
4. Wilayah Operasi
5. Nomor Plat
6. Kondisi Bus
7. Kapasitas Penumpang
8. Kode Trayek

Variabel atau fitur-fitur tersebut terdiri dari fitur bertipe data integer yaitu Tahun Pengadaan. Sedangkan fitur lainnya bertipe data objek atau kategori.

Kedua data tersebut dapat dihubungkan untuk mengetahui faktor lain yang mungkin saja bisa mempengaruhi banyaknya penumpang armada bus sekolah. Pada data jumlah armada, terdapat fitur Kondisi Bus yang terdiri dari "Baik", "Kurang Baik" dan "Rusak". Selain itu terdapat fitur **Kode Trayek** yang menunjukkan kode dari bus sekolah yang beroperasi di Zona tertentu. Sedangkan pada Data Penumpang Bus, kita telah mengetahui Zona mana saja yang ada dan jumlah penumpangnya. Apabila dapat menghubungkan antara **Kode Trayek** dengan **Daerah Operasi (Zona)** kita bisa melihat apakah faktor kondisi bus mempengaruhi banyaknya penumpang di suatu daerah. Namun, dalam hal ini belum terdapat data pasti mengenai kode trayek apa yang beroperasi di zona tertentu. Belum terdapat data pendukung yang lengkap. Akan tetapi, kedua fitur tersebut dapat dihubungkan, karena berdasarkan *timeline* akun Twitter [@bussekolah_dki](#) bahwa contohnya saja trayek Z32 beroperasi di Zona 9. Maka perlu adanya penambahan data apabila ingin mengetahui faktor-faktor lain terkait jumlah penumpang terhadap armada bus sekolah di daerah tertentu.

3. Data Preparation

Pada data pertama yaitu Data Penumpang Bus Sekolah Tahun 2017 terdapat beberapa data missing value pada fitur jumlah bus dan jumlah penumpang di bulan dan daerah operasi tertentu. Missing value tersebut diisi dengan nilai rata-rata dari fitur.

Terdapat dua data yang dapat dihubungkan seperti yang telah dijelaskan pada Data Understanding. Akan tetapi data tersebut belum lengkap, dan perlu dilakukan *enrichment* dan penambahan data agar dapat saling dihubungkan. Penambahan data yang diperlukan yaitu perlu diketahui kode trayek bus sekolah beroperasi di daerah operasi mana. Seperti kode trayek Z32 beroperasi di daerah Zona 9, Z45 untuk daerah Zonasi 11, Z25 untuk daerah Zonasi 10 dan lain-lain. Beberapa yang disebutkan tersebut didapatkan melalui *timeline* Twitter akun resmi Bus Sekolah DKI. Namun, belum didapatkan data lengkap mengenai kode trayek dengan daerah operasinya.

4. Analysis and Modelling

Dalam menyelesaikan problem ini digunakan model analisis deskriptif. Analisis deskriptif digunakan karena dengan membandingkan beberapa variabel saja tanpa perlu menggunakan model prediksi sudah dapat dilihat daerah mana yang membutuhkan sosialisasi kepada para penumpang dan daerah yang membutuhkan penambahan armada. Data yang digunakan adalah gabungan dari semua Data Penumpang Bus Sekolah Tahun 2017 dari setiap bulannya.

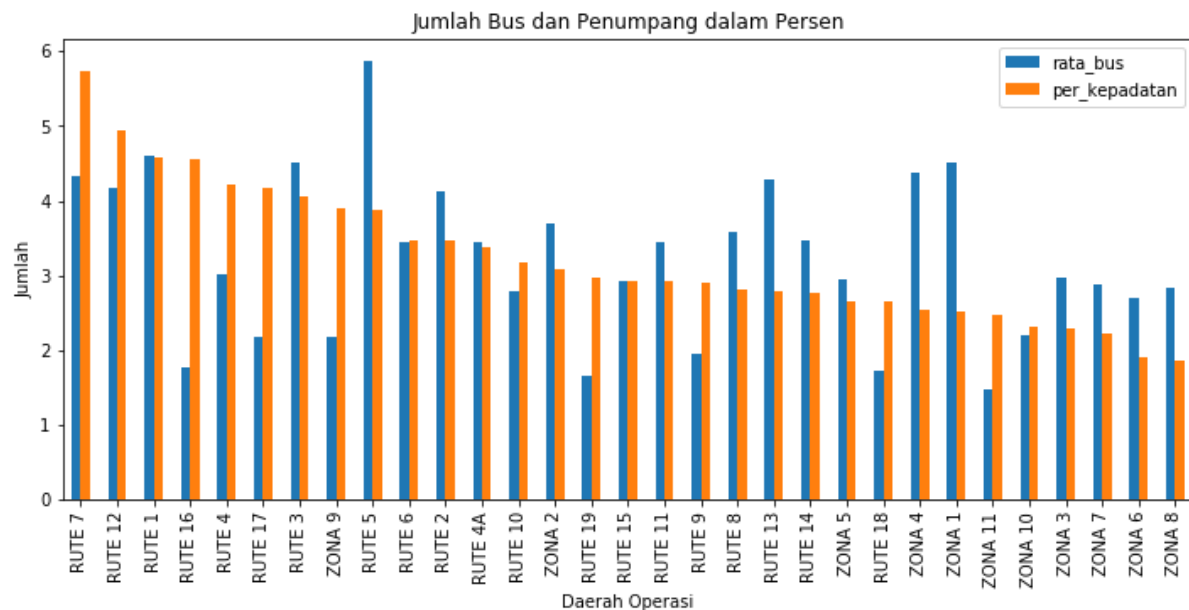
5. Evaluation

Pada tahap evaluation tidak digunakan metode testing atau confusion matrix seperti pada pemodelan prediktif.

6. Visualization and Presentation

Visualisasi di bawah ini merupakan plot perbandingan jumlah rata-rata penumpang dan banyaknya bus di setiap daerah operasi di tahun 2017. Visualisasi didapatkan berdasarkan data pertama. Dapat dilihat bahwa terdapat rute yang jumlah kepadatan penumpangnya tidak sebanding dengan armada yang disediakan. Misalnya pada Rute 7, bahwa penumpang belum semua bisa terpenuhi oleh armada yang disediakan.

Sedangkan pada daerah seperti Rute 5, terdapat banyak armada yang belum bisa dimaksimalkan oleh penumpang karena penumpang yang masih sedikit.



Sedikitnya penumpang tersebut perlu juga diketahui apakah terdapat pengaruh faktor-faktor lain. Untuk itu perlu dilakukan iterasi pada proses Data Understanding dan Preparation untuk menambahkan data yang dapat dihubungkan dengan variabel lain seperti variabel kondisi bus sekolah yang dibahas pada data preparation. Setelah dilakukan iterasi berulang kali hingga ditemukan faktor yang tepat mengapa adanya ketimpangan antara jumlah armada yang ada dan penumpang, maka dapat digunakan dalam memutuskan strategi pengadaan atau peningkatan fasilitas armada oleh pemerintah DKI Jakarta.