



Minggu ke-16

## Final Test

### Topik: Clustering for Prediction

Ali Ridho Barakbah

Knowledge Engineering Research Group  
Department of Information and Computer Engineering  
Politeknik Elektronika Negeri Surabaya

# Transaction Dataset

InvoiceNo	StockCode	Qty	InvoiceDate	CustomerID	Country
537626	22725	830	12/7/10 14:57	12347	Iceland
537626	22729	948	12/7/10 14:57	12347	Iceland
537626	22195	695	12/7/10 14:57	12347	Iceland
542237	22725	636	1/26/11 14:30	12347	Iceland
542237	22729	536	1/26/11 14:30	12347	Iceland
542237	47559	919	1/26/11 14:30	12347	Iceland
542237	21154	803	1/26/11 14:30	12347	Iceland
542237	21035	532	1/26/11 14:30	12347	Iceland
...	...		...	...	...

# Assignment #

1. **dataset**  $\leftarrow$  transaction.csv, dan tampilkan
2. **data**  $\leftarrow$  ambillah data pada **dataset** untuk fitur InvoiceNo, Country, month, year ("2011")

**Lakukan secara iteratif dari month, dimana  $i=1-12$**

3. **data\_bulan**  $\leftarrow$  ambillah **data** pada bulan ke- $i$
4. **transaksi**  $\leftarrow$  hitunglah banyaknya transaksi pada **data\_bulan** untuk setiap negara (1 kode InvoiceNo = 1 transaksi)
5. **cluster**  $\leftarrow$  lakukan clustering pada **transaksi** dengan Average Linkage, dengan  $k=3$
6. **centroid**  $\leftarrow$  tentukan posisi centroid dari setiap **cluster**
7. **sorted**  $\leftarrow$  lakukan pengurutan posisi **centroid** secara ascending
8. **transaksi\_tinggi[i]**  $\leftarrow$  ambil nilai **sorted** pada indeks terakhir, yang mengindikasikan cluster dari negara-negara yang tergolong transaksi tinggi

9. Visualisasikan pergerakan nilai **transaksi\_tinggi** dimana sumbu  $x=\text{month}(1-12)$  dan sumbu  $y=\text{transaksi_tinggi}$
10. PredictedVal  $\leftarrow$  prediksi nilai transaksi tinggi dari **transaksi\_tinggi** untuk bulan Januari 2012 dengan Linear Regression

# Pengumpulan Tugas

---

- Buatlah coding dengan Bahasa pemrograman/tools apapun untuk semua assignment
- Buatlah laporan dalam slide ppt. Laporan terdiri dari screenshot coding dan hasil running untuk setiap assignment.
- Simpan laporan dalam file pdf dengan format penamaan:  
DM\_UAS\_NRP\_namadepan.pdf
- Upload file tersebut di alamat ini:  
<http://ridho.lecturer.pens.ac.id/lecture.php>  
(Pilih Mata Kuliah : Data Mining)
- Deadline upload: Jumat, 25 Juni 2021