

## **BAB II**

### **LANDASAN TEORI**

#### 2.1. Visualisasi Data

##### 2.1.1 Data Visualisasi

Mengubah atau mentransformasi data mentah menjadi suatu bentuk grafik, dengan tujuan manusia dapat mudah menerima informasi dari data tersebut sesuai dengan persepsinya masing-masing, disebut dengan data visualisasi[4]. Data visualisasi juga dapat didefinisikan sebagai suatu cara untuk melakukan presentasi informasi dalam bentuk grafik, dapat memungkinkan para pembaca melihat suatu pola atau tren dari hasil transformasi data tersebut. Data visualisasi juga dapat digunakan untuk melakukan eksplorasi data dengan menggunakan transformasi data menjadi grafik, warna, ukuran, bentuk, yang mewakili setiap aspek dari kumpulan data yang sedang dianalisa[5]. Hasil visualisasi data seperti statistik, diagram, grafik merupakan hasil dari penyederhanaan data[6].

Data visualisasi data dapat digunakan untuk memberikan gambaran secara umum terkait data yang banyak jumlahnya, serta membantu ilmuwan untuk melakukan analisis data[2]. Data visualisasi yang efektif harus mudah dipahami oleh seorang ilmuwan, dengan tujuan ilmuwan dapat menginformasikan hasil data visualisasinya, data visualisasi tidak terpisahkan dari suatu proses ilmiah[7].

##### 2.1.2 Visual Analitik

Visual analitik merupakan perkembangan dari visualisasi informasi. Visual analitik merupakan “metafora visual yang dikombinasikan dengan wacana atau informasi dari manusia, yang memungkinkannya menemukan suatu informasi yang tak terduga dari hasil data visualisasi tersebut”. Data

analisis dan data visualisasi dapat dikombinasikan sebagai media dan juga interaksi untuk proses analisis data. Visual analitik ditujukan untuk merubah suatu informasi yang berlebihan. Informasi yang berlebihan tersebut digunakan menjadi peluang sebagai saran penilaian[8].

Ada beberapa definisi dari visual analitik, seperti sebagai alat untuk pendekatan yang terintegrasi dari menggabungkan data analisis dengan visualisasi. Serta visual analitik dapat dikatakan sebagai metode untuk melakukan analisis terhadap data untuk mendapatkan wawasan dan pengetahuan yang tak terduga dari data mentah tersebut. Visualisasi analitik dengan menggunakan visualisasi dapat mempresentasikan suatu data yang dimana manusia dapat melihat hasil visualisasi data tersebut dengan persepsi yang berbeda. Tidak hanya sekedar mempresentasikan data saja, berdasarkan hasil visualisasi tersebut dapat membantu dalam data analisis[8].

## 2.2 *Exploratory Data Analysis*

### 2.2.1 Pengertian *Exploratory Data Analysis*

*Exploratory Data Analysis* adalah suatu metode eksplorasi terhadap suatu data. *Exploratory Data Analysis* dapat menggunakan teknik aritmatika ataupun teknik visual grafik. Dari teknis tersebut dapat membuat hasil rangkuman atau pengolahan dari data mentah yang ada. *Exploratory Data Analysis* ini dapat digunakan untuk mencari suatu pola atau tren yang ada dari hasil visualisasi data[9].

### 2.2.2 Sejarah *Exploratory Data Analysis*

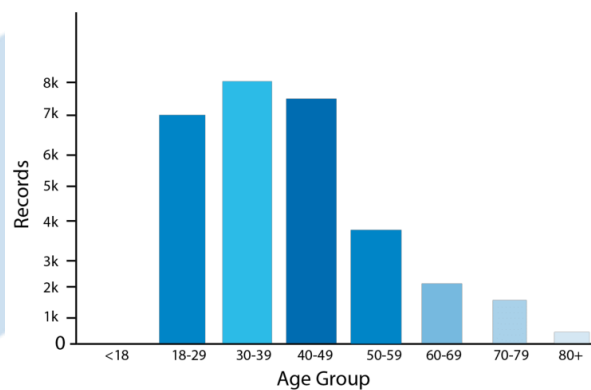
*Exploratory Data Analysis* awalnya dikenalkan oleh John Tukey dan rekannya sebagai ahli statistik, pada saat itu John Tukey berpendapat “terlalu sedikit penggunaan data sebagai sarana pengembangan teori”. Terkait segala pertanyaan pada penelitian *Exploratory Data Analysis* mengartikulasikannya dengan baik, dapat menemukan pola atau fenomena

yang tidak terduga, dan hipotesis yang kuat. Ada beberapa prinsip yang mendasari *Exploratory Data Analysis*, seperti fleksibel saat digunakan untuk data data tertentu, visualisasi, pembuatan model, dan lain-lain[10].

Visualisasi data yang digunakan dapat berupa *diagram*, *bar*, *plot*, dan lain-lainnya. Dengan adanya *exploratory data analysis*, data yang banyak dan mentah tersebut memungkinkan mendapatkan informasi. Dapat lebih mudah melihat informasi dari hasil *exploratory data analysis*[9]. *Exploratory Data Analysis* banyak dipergunakan untuk melakukan analisis terkait berbagai aspek. Diagram kotak yang terdapat pada *Exploratory Data Analysis* memberikan informasi terkait jangkauan dan ruang lingkup, *exploratory data analysis* tidak menggunakan dimensi spasial dan juga lokasi objek pengamatan[11].

#### 2.2.2.1 Bar Chart

Penggunaan *bar chart* untuk melakukan perbandingan keseluruhan dari data , dan juga dapat menampilkan data tertinggi dan data terkecil[12].



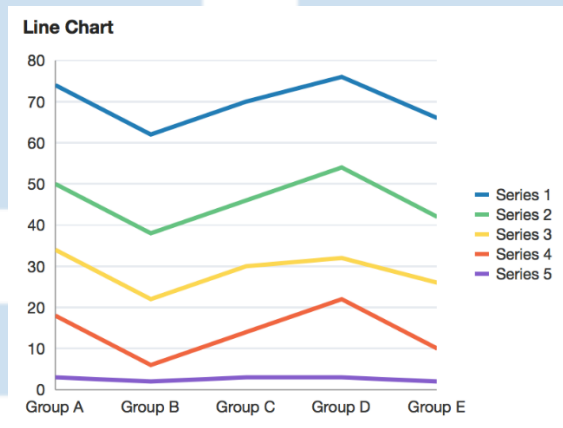
Gambar 2. 1 Barchart

#### 2.2.2.2 Line Chart

*Line chart* menghubungkan dari titik ke titik lainnya.

Penggunaan line chart bertujuan untuk dapat melihat suatu tren dari

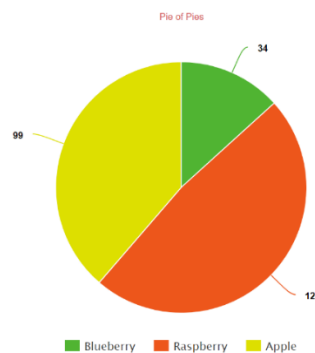
waktu ke waktu, serta dapat digunakan untuk melakukan perkiraan suatu tren atau nilai di masa mendatang[12].



Gambar 2.2 Linechart

### 2.2.2.3 Pie Chart

Penggunaan pie chart untuk menampilkan proporsi dari keseluruhan data yang ada. *Pie chart* dapat digunakan untuk menambahkan detail dari visualisasi yang ada[12].



Gambar 2.3 Piechart

## 2.3 Python

Python merupakan salah satu bahasa pemrograman yang cukup umum dan sering digunakan untuk penelitian *data science*. Python tidak hanya digunakan untuk *data science*, bisa juga digunakan untuk *network automation*. Tetapi Python

memiliki alat atau *library* yang kuat atau kompeten untuk melakukan data analisis. Beberapa contoh *library* yang digunakan untuk penerapan data analisis adalah *library* Pandas, Numpy, Matplotlib. Untuk Pandas merupakan *library* pada Python yang digunakan untuk melakukan pemrosesan pada data. Untuk *library* Numpy memiliki fungsi untuk menghitung numerik pada data berbasis *array*. Untuk *library* matplotlib digunakan untuk pembuatan grafik, *library* tersebut sering digunakan untuk metode *exploratory data analysis*[13].

Alasan kenapa Python biasa digunakan untuk melakukan data analisis adalah memiliki banyak *library* untuk statistik, dan juga analisis numerik. Bahasa pemrograman Python memiliki banyak *code* yang bisa diakses secara gratis[14]. Python dapat mengolah data yang kapasitasnya besar, dan juga dapat model *code* yang telah dibangun dapat dipindahkan di *platform* yang lain. Dikarenakan Python memiliki banyak *library* yang mendukung untuk pengolahan data, dapat menciptakan hasil visualisasi data dengan jelas dan mudah dipahami[15].

#### 2.4 Simple Moving Average

*Simple Moving Average* merupakan salah satu metode untuk melakukan peramalan terhadap suatu data. *Simple Moving Average* melakukan peramalan terhadap suatu data dengan melakukan perhitungan rata-rata yang ada pada rentang waktu tertentu. Algoritma *Simple Moving Average* baik digunakan pada tipe data yang tidak memiliki tren, serta tidak menggunakan pembobotan[16]. Dengan menggunakan data historis untuk menghasilkan suatu prediksi yang baik pada data yang tidak memiliki pola.

$$Y_{t-1} = Y_t + Y_{t-1} + \dots + Y_{t-m+1}$$

Persamaan 2.1 Simple Moving Average

Pada persamaan 2.1  $Y_t$  merupakan data yang ada pada periode saat ini, atau yang digunakan, untuk  $m$  merupakan periode yang digunakan untuk melakukan

prediksi terhadap data yang digunakan. Data yang ada pada Y akan digabungkan untuk mencari suatu nilai prediksi[16].

### 2.5 Chi Square

*Chi square* merupakan suatu algoritma yang digunakan untuk melakukan pengujian terhadap dua variabel. *Chi square* juga merupakan tes numerik untuk melakukan pengukuran keterkaitan antara dua variabel. Melakukan pengujian dengan *chi square* perlu mengikuti syarat-syarat yang ada seperti tidak adanya sel dengan nilai nyata atau bisa juga dengan *actual count* sebesar 0, apabila tabel yang digunakan 2 x 2 maka dalam 1 sel tidak boleh memiliki nilai *expected count* kurang dari 5.

$$\chi^2(t_k, c_j) = \frac{N \cdot (A \cdot D - B \cdot C)^2}{(A+B) \cdot (C+D) \cdot (A+C) \cdot (B+D)}$$

Persamaan 2.2 Chi Square

Pada persamaan 2.2 N menunjukkan banyaknya data, A menunjukkan probabilitas yang mungkin terjadi pada data independent, B menunjukkan kemungkinan yang terjadi pada data dependent, C menunjukkan probabilitas yang tidak dimiliki oleh data independent, D menunjukkan probabilitas pada data independent dan bukan di dalam data independent. Pada persamaan 2.1 digunakan untuk melakukan pemilihan terhadap atribut atau variabel yang ada pada dataset, sehingga dapat membantu mengurangi rentang yang ada pada atribut atau variabel yang tidak sesuai[17].

### 2.6 Skala Likert

Skala Likert adalah skala psikometrik yang digunakan pada penelitian dalam mengukur pendapat seseorang. Skala likert digunakan untuk mengukur tingkat persetujuan dalam suatu pertanyaan yang telah diberikan pada sebuah kuesioner. Skala Likert juga dapat digunakan dalam mengukur data kualitatif maupun kuantitatif. Skala Likert memiliki berbagai macam bentuk, sehingga dapat

disesuaikan dengan tujuan yang ingin dicapai oleh peneliti. Salah satu contoh bentuk skala Likert untuk mengukur persetujuan terhadap pernyataan memiliki lima pilihan, seperti sangat setuju, setuju, netral, tidak setuju, sangat tidak setuju. Dari pilihan yang diberikan penilaian 5,4,3,2,1.

$$T x P_n$$

Persamaan 2.3 Skala Likert

Pada persamaan 2.2 merupakan rumus untuk menghitung hasil dari kuesioner yang dijawab atau diisi oleh responden. T merupakan total responden yang memilih, lalu untuk P<sub>n</sub> merupakan pilihan angka dari skor skala likert [18].

## 2.7 Penelitian Terdahulu

Tabel 2.1 Penelitian Terdahulu

Judul Artikel	Nama Jurnal	Penulis	Hasil	Kesimpulan
Analisis Toko dan Asal Toko Fashion Pria di Shopee Menggunakan Data Scraping dan Exploratory Data Analysis	Majalah Ilmiah Teknologi Elektro, Vol.21, 2022	M. E.-K. Kesuma and R. Iskandar	Menggunakan metode EDA (Exploratory Data Analysis) untuk melakukan analisa terkait fashion pria di shopee, dapat memberikan hasil pengolahan data berupa grafis yang dapat membantu peneliti dalam menjawab pertanyaan, antara lain, Apa nama toko yang penjualan terbanyak di pasar kategori fashion pria?, Kota asal toko yang penjualan terbanyak di pasar kategori fashion pria?.	Dengan menggunakan metode Exploratory Data Analysis dapat mengetahui toko yang sedang tren, kota asal berdasarkan produk yang paling banyak dibeli. Pada peringkat pertama terdapat toko erigo yang telah melakukan penjualan produk lebih dari 400.000. Untuk produk yang paling banyak dibeli di aplikasi shopee, terdapat baju.

Tor Marketplaces Exploratory Data Analysis : The Drug Case	Communications in Computer and Information Science, Vol.2, 2021	Jahankhani, Hamid Carlile, Alex Emm, David Hosseinian-Far, Amin Brown, Guy Sexton, Graham Jamal, Arshad	Penggunaan EDA (Exploratory Data Analysis) pada penelitian ini adalah untuk dapat mengekstrak data dan menemukan variabel yang penting untuk penelitian ini. Untuk menghindari adanya penipuan tersbut, peneliti menggunakan vendor yang penawarannya lebih dari 3. Dengan menggunakan EDA peneliti dapat menemukan zat yang paling diminati oleh customer, yaitu zat east india. Dapat menemukan zat tersebut paling banyak dijual di benua mana, yaitu eropa.	Situs perdagangan pada Darknet, memiliki peranan yang cukup besar sebagai pemicu kejahatan. Para penjual dapat melakukan perdagangan zat illegal tersebut dengan menggunakan jaringan TOR secara anonim. Karena dampak dari perdagangan zat illegal menggunakan jaringan TOR masih belum jelas. Menggunakan EDA merupakan landasan yang penting untuk memantau pola dan tren yang ada.
A Methodology for Securities and Cryptocurrency Trading Using Exploratory Data Analysis and Artificial Intelligence	2021 1 <sup>st</sup> International Conference on Artificial Intelligence and Data Analytics, CAIDA 2021, 2021	Al-Ameer, Ali Al-Sunni, Fouad	Dengan menggunakan EDA (Exploratory Data Analysis) untuk sekuritas dan perdagangan dalam cryptocurrency, mendapatkan pola terkait perdagangan cryptocurrency. Dengan hal tersebut dapat mengetahui kapan para investor untuk menjual cryptocurrency nya. Dengan metode EDA dapat mengetahui data	Menggunakan metode EDA dapat menentukan model yang tepat untuk membangun model artificial intelligence pada penelitian sekuritas dan perdagangan cryptocurrency. Untuk model naïve bayes classifier tidak dapat menangkap data yang kompleks.



			<p>mana yang memiliki pengaruh paling besar terkait perdagangan pada cryptocurrency untuk dibangun artificial intelligence. Algoritma Naïve Bayes Classifier memiliki hasil akurasi paling rendah untuk sekuritas dan perdagangan pada cryptocurrency.</p>	
Exploratory Data Analysis As A Foundation Inductive Research	Human Resource Management Review, Vol.27, 2017	Jebb, Andrew T. Parrigon, Scott Woo, Sang Eun	<p>Penggunaan Exploratory Data Analysis dalam melakukan analisa terhadap penemuan dan pendeteksian empiris pada data. Penggunaan Exploratory Data Analysis memberikan dampak dalam memaksimalkan nilai yang terdapat pada data. Exploratory Data Analysis memiliki banyak teknik dalam penggunaan kuantitatif untuk mencari suatu pola atau struktur pada data.</p>	<p>Menggunakan metode Exploratory Data Analysis dapat membantu sains dalam melengkapi antara eksplorasi data dan konfirmasi data. Eksplorasi data pada tahapan Exploratory Data Analysis memberikan dampak dalam landasan pada ide awal dalam membangun analisa data.</p>

Pada Tabel 2.1 jurnal pertama yang berjudul “Analisis Toko dan Asal Toko Fashion Pria di Shopee Menggunakan Data Scraping dan Exploratory Data Analysis” mendapatkan kesimpulan toko mana yang memiliki penjualan paling banyak untuk kategori fashion pria, dan juga dapat melihat pakaian jenis apa yang paling diminati oleh pria. Pada jurnal kedua yang berjudul “Tor Marketplaces

*Exploratory Data Analysis : The Drug Case*” mendapatkan tren dan pola penjualan zat ilegal menggunakan jaringan *TOR*, serta mengetahui zat ilegal apa yang paling diminati oleh *customer*. Serta dapat juga mengetahui benua mana yang paling banyak melakukan transaksi jual beli zat ilegal. Pada jurnal ketiga “*A Methodology for Securities and Cryptocurrency Trading Using Exploratory Data Analysis and Artificial Intelligence*” mendapatkan bahwa dengan *exploratory data analysis* dapat menentukan variabel yang penting untuk mendukung pembuatan model *artificial intelligence* dalam penelitian sekuritas dan perdagangan *cryptocurrency*.

Berdasarkan penelitian terdahulu, seperti yang telah dicantumkan pada tabel 2.1, bahwa metode *exploratory data analysis* (EDA) dapat membantu para peneliti, perusahaan dan lain-lain, untuk mengetahui suatu pola dan tren yang ada pada data. Implementasi Visual Analitik menggunakan metode *exploratory data analysis* pada perusahaan dapat membantu mengolah data mentah, menjadi data visual yang mudah dipahami oleh manusia. Pertama mencari variabel yang penting pada data untuk divisualisasikan, dapat memberikan informasi bagi perusahaan. Serta mendapatkan informasi baru dari hasil pengolahan data tersebut. Dapat membantu perusahaan dalam menentukan strategi bisnis dan juga pengambilan keputusan. Pada penelitian ini akan di bangun visualisasi analitik pada data penjualan PT. Wahana Internet Nusantara. Data penjualan yang digunakan data penjualan tahun 2022. Visualisasi analitik data akan dibangun menggunakan metode *exploratory data analysis*. Hasil dari visualisasi data yang didapat dengan menggunakan metode *exploratory data analysis*, akan ditampilkan didalam *prototype* suatu *website*.