

BAB I

PENDAHULUAN

1.1 Latar Belakang

Belakangan ini perkembangan Teknologi Informasi (TI) telah semakin meningkat. Hal ini menyebabkan persebaran data yang semakin pesat, sehingga untuk menemukan data di era digital ini tidaklah begitu sulit lewat internet. Banyaknya dataset terbuka yang bebas diakses berdampak pada informasi berharga yang bisa didapatkan dan dimanfaatkan untuk keperluan bisnis, pemerintahan, ataupun pribadi (Syaripul & Bachtiar, 2016). Namun untuk mengubah data mentah tersebut menjadi sebuah informasi berharga perlu dilakukan proses analisis pada data tersebut. Dengan adanya beberapa dataset yang sangat besar, kompleks, dan sulit dianalisis, akan memerlukan sebuah bantuan alat (*tools*). Salah satu metode yang paling membantu dalam memahami data adalah visualisasi data (Microsoft, 2018). Sebuah studi menunjukkan bahwa manusia merespons visual lebih baik daripada rangsangan jenis lain. Otak manusia memproses informasi visual 60.000 kali lebih cepat daripada teks. Faktanya, data visual menyumbang 90 persen dari informasi yang dikirimkan ke otak (Oracle, 2017).

Salah satu metode statistika yang biasa digunakan untuk menganalisis data yaitu metode statistika deskriptif (*descriptive statistic*). Statistika deskriptif merupakan statistika yang digunakan untuk menggambarkan atau menganalisis suatu statistik hasil penelitian, tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas. Walaupun analisis data dengan menerapkan metode statistika deskriptif ini dinyatakan sebagai analisis sederhana, akan tetapi hasil analisis tersebut dapat menjadi masukan yang sangat berharga untuk para pengambil keputusan, tergantung pada bentuk dan cara menyajikan hasil analisis tersebut (Agung, 2019).

Banyak aplikasi yang bisa digunakan untuk melakukan analisis data, baik aplikasi web maupun dekstop. Salah satu contoh dalam bentuk *web* yaitu *website* *datawrapper.de* (2018). *Website* tersebut bisa dimanfaatkan untuk melakukan

Aksan Maulana, 2019

SISTEM ANALISIS DATA MENGGUNAKAN DESCRIPTIVE STATISTIC BERBASIS WEB

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

analisis data dikarenakan penyajian data pada website tersebut sudah dalam bentuk visual yaitu dalam bentuk *chart*, *diagram* maupun *graph*. Namun *website* tersebut merupakan hasil karya orang lain, sehingga kita tidak bisa melakukan pengembangan pada aplikasi itu lebih mendalam dikarenakan *source code* yang tidak dimiliki.

Oleh karena itu, dalam penelitian ini akan membangun sebuah sistem yang serupa dengan *website* *datawrapper.de* untuk membantu proses analisis data menggunakan metode *descriptive statistic* dengan penyajian data dalam beberapa bentuk visual yang berbasis web dengan *source code* yang terbuka sehingga bisa dikembangkan kembali pada penelitian selanjutnya. Banyak bahasa pemrograman yang bisa digunakan untuk melakukan *data analysis*, salah satunya adalah bahasa pemrograman *Python*. *Python* sangat mendukung dalam kegiatan analisis dan *mining* data sehingga banyak para pengembang yang menggunakan atau bahkan mengembangkan *library* yang bisa digunakan untuk analisis data. Beberapa *library* yang bisa digunakan dalam bidang *data scientis* yaitu *Numpy*, *Scipy*, *Matplotlib*, *Pandas* dan *Scikit-learn*.

1.2 Rumusan Masalah

Berdasarkan permasalahan yang muncul dalam latar belakang, dapat disimpulkan beberapa rumusan masalah dalam penelitian ini:

1. Bagaimana membangun sistem yang menampilkan data mentah menjadi bentuk visual yang memberikan kemudahan untuk mengambil informasi pada data tersebut?
2. Bagaimanakah bentuk struktur tabel dari data mentah yang bisa memudahkan proses penyajian data menjadi bentuk visual?
3. Apakah setiap data yang disajikan dengan berbagai bentuk visual mampu memberikan hasil optimal?

1.3 Tujuan Penelitian

Adapun tujuan dalam penelitian ini adalah sebagai berikut:

1. Membangun sistem analisis data berbasis web menggunakan metode *descriptive statistic* dengan penyajian data dalam beberapa bentuk visual.
2. Menentukan bentuk data dengan struktur tabel yang mudah untuk divisualisasikan.
3. Mengecek keoptimalan untuk setiap bentuk visual terhadap data.
4. Evaluasi dan uji coba fungsional dari sistem yang dibangun.

1.4 Batasan Masalah

Adapula batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Ekstensi file yang digunakan yaitu .xlsx, .xls, .txt dan .csv saja.
2. Struktur tabel harus dengan ketentuan seperti: baris pertama merupakan *title* sedangkan baris kedua sampai R rows merupakan *value*; kolom pertama dianjurkan bertipe *date/datetime* karena akan menjadi label untuk sumbu X atau istilah lain menjadi *series*, sedangkan sisa *value* kolom berikutnya sampai C column harus bersifat *numerik*.
3. Bahasa pemrograman yang digunakan adalah *Django Web Framework Python*.