

PROPOSAL PROYEK
12S4056 – Data Visualization

Video *Games* Sales Visualization



Disusun oleh:

12S17008 Rezky Simanjuntak
12S17060 Siti Berliana Manurung
12S17061 Mega Christy Silalahi

PROGRAM STUDI SARJANA SISTEM INFORMASI
FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO
INSTITUT TEKNOLOGI DEL
NOVEMBER 2020

DAFTAR ISI

DAFTAR ISI.....	ii
1. Pendahuluan	3
1.1 Latar Belakang	3
1.2 Tujuan.....	4
1.3 Manfaat.....	4
1.4 Ruang Lingkup Proyek	4
2. Bab Isi	5
2.1 Aktifitas Pengembangan Dashboard Visualisasi Data	5
3. Rencana Kerja.....	9
3.1 Jadwal Kegiatan	9
3.2 Pembagian Tugas	9
4. Referensi	10

DAFTAR TABEL

Tabel 1 Tipe Atribut Pada Video <i>Game Sales</i>	6
Tabel 2 Jadwal Kegiatan Proyek.....	9
Tabel 3 Pembagian Tugas.....	9

DAFTAR GAMBAR

Gambar 1 Aktivitas Pengembangan Dashboard Visualisasi Data	5
Gambar 2 Tipe Evaluasi	7

1. Pendahuluan

Pada bab ini berisi penjelasan mengenai latar belakang pengerjaan proyek, tujuan, manfaat, dan ruang lingkup dalam proyek.

1.1 Latar Belakang

Dalam dunia teknologi, video *games* sudah berkembang pesat dimana sudah menyediakan berbagai variasi *game* baik dari genre *game* dan ragam usia mulai dari anak kecil hingga dewasa. Hal ini menyebabkan lonjakan penjualan video *game* saat ini. Video *game* dirilis oleh penerbit besar di banyak platform perangkat keras populer. Ini memberikan satu-satunya pengalaman nyata dari hiburan interaktif yang ditawarkan oleh teknologi modern. Ini juga menyediakan bentuk hiburan yang berkembang pesat dan digunakan untuk tujuan pendidikan serta bisnis.

Dalam dekade sebelumnya, beberapa rilis video *game* besar telah meningkatkan standar barang hiburan konvensional dalam hal pendapatan yang diperoleh. Ada beberapa jenis video *game* blockbuster yang dijual hari ini seperti “*Grand Theft Auto V*” oleh *Rock Star Games* dan “*Call of Duty Series*” oleh *Activision*. Jenis video *game* ini telah menghasilkan serangkaian rekor pendapatan tahunan selama periode tiga tahun. Pemodelan prediktif telah lama menjadi tujuan banyak individu dan organisasi. Sains ini memiliki banyak teknik, dengan simulasi dan pembelajaran mesin sebagai intinya. Selain potensi simulasi, teknik pembelajaran mesin dikenal karena kemampuannya dalam mengungkap tren data yang tersembunyi. Meskipun pilihan algoritma yang digunakan dalam setiap studi mungkin berbeda, semuanya memiliki satu kesamaan yang sama, mereka memberikan pilihan yang dapat menggunakan data untuk menciptakan peluang dalam menarik keuntungan dari penjualan oleh Penerbit. Penerbit dan pengembang ini berada dalam bisnis menghasilkan uang dan menghasilkan penjualan, jadi mereka ingin melihat variabel mana yang paling penting untuk penjualan video *game* mereka. Idealnya, hasil analisis visualisasi ini diharapkan akan dapat membantu pengembang dalam memutuskan di *platform*, *genre*, atau negara mana yang akan merilis video *game* masa depan mereka. Video *game* tidak lagi menjadi produk khusus dengan segmen konsumen yang kecil. Industri ini telah memengaruhi sebagian besar segmen konsumen, di mana sebagian besar populasi saat ini telah tumbuh dengan video *game* dan memainkannya untuk hiburan atau pendidikan dengan cara tertentu. Selain itu, transformasi pohon klasifikasi menjadi model analitik keputusan juga telah digunakan dalam menyelesaikan kebijakan pengembangan *game* yang memaksimalkan nilai. Adapun berdasarkan hal tersebut maka maksud dari proyek ini adalah untuk mengetahui faktor-faktor yang menyebabkan penjualan video *game* menjadi populer dan sukses di bidang finansial. Adapun data yang dikumpulkan terdiri dari sekitar 16.598 data yang berkaitan dengan judul video *game* individu yang dirilis di berbagai platform. Proyek ini menggunakan kumpulan data unik judul video *game* untuk memperkirakan pengaruh variabel dependen yang digunakan untuk secara empiris penjualan unit *game* ke berbagai negara mulai tahun 1980 hingga 2020, yang diperoleh dari *dataset online* yang dikelola oleh *Kaggle.com*. Selain menemukan video *game* menjadi populer berdasarkan jumlah penjualan *game*, visualisasi juga menampilkan Publisher dengan *game* yang paling laris.

1.2 Tujuan

Adapun Tujuan dari proyek ini adalah sebagai berikut :

1. Memvisualisasikan Video *Game* terpopuler dalam 10 tahun terakhir berdasarkan *Nama Game*.
2. Memvisualisasikan Video *Game* terpopuler dalam 10 tahun terakhir berdasarkan hasil produksi perusahaan (*Publisher*) Nintendo.
3. Menentukan penjualan Video *Games* terlaris dalam 5 tahun terakhir di negara Jepang.

1.3 Manfaat

Visualisasi data dilakukan untuk merepresentasikan grafik pada data yang dapat membantu pembaca dalam memahami makna dari data dengan cepat. Dalam melakukan visualiasi penjualan video *games* diharapkan dapat memberikan kemudahan bagi pihak perusahaan *game* dan *game store* dalam melakukan manajemen dan pengambilan keputusan untuk menentukan keputusan atau pilihan yang terbaik. Dengan adanya visualisasi tersebut maka perusahaan *game* dapat menentukan video *games* apa yang paling banyak disukai atau digunakan masyarakat, sehingga dapat meningkatkan produksi video *games*. Begitu juga bagi *game store* yang akan dapat melihat penjualan video *games* yang paling laris atau video *game* yang paling sering dibeli oleh masyarakat, dari hasil visualiasi tersebut maka para *game store* akan lebih mudah menentukan video *games* apa yang akan dijual selanjutnya. Visualisasi ini juga dapat memberikan informasi kepada pengguna *game* atau *gamers* untuk melihat informasi *game* apa yang paling populer.

1.4 Ruang Lingkup Proyek

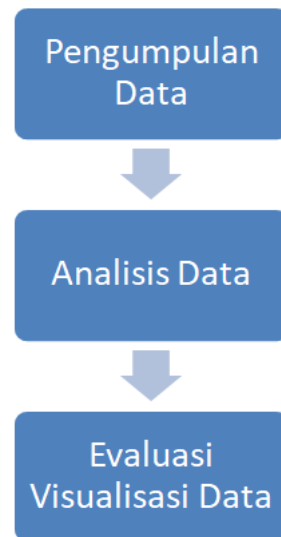
Adapun beberapa hal yang digunakan sebagai ruang lingkup dalam proyek yaitu:

1. Ketersediaan *dataset* (*dataset availability*) pada dataset Video *Game Sales* adalah *dynamic*, yang berarti dataset tersebut dapat diakses melalui online.
2. *Tools* yang akan digunakan dalam pengerjaan proyek ini untuk memvisualisasikan data tersebut adalah *Tableau*.
3. Bahasa yang digunakan dalam *dataset* Video *Game Sales* menggunakan Bahasa Indonesia.
4. Dataset Video *Game Sales* menampilkan penjualan video *game* untuk negara jepang dan benua Eropa dan Amerika Utara.

2. Bab Isi

Pada bab ini akan menjelaskan mengenai aktifitas pengembangan *dashboard* visualisasi data yang dibuat dalam bentuk diagram alir.

2.1 Aktifitas Pengembangan Dashboard Visualisasi Data



Gambar 1 Aktifitas Pengembangan Dashboard Visualisasi Data

1. Pengumpulan Data

Dataset yang akan digunakan dalam visualiasasi data adalah data *Video Games Sales*, yang dipublikasikan oleh *Gregory Smith* pada *Kaggle* yang dapat diakses di <https://www.kaggle.com/regorut/videogamesales>. Data *Video Games Sales* dikumpulkan dari berbagai *publisher* atau perusahaan yang memproduksi *Video Games* untuk setiap tahunnya. Data tersebut juga menyediakan jenis *platform* yang paling sering digunakan dalam *video games*, tahun perilisan *video game*, *genre game*, dan juga penjualan *video games* ditampilkan untuk benua Amerika Utara, Eropa, dan untuk negara Jepang. Untuk penjualan lainnya dikalkulasikan dalam *Others_Sales* dan total penjualan secara global.

2. Analisis Data

a. Tipe data (*data type*)

Tipe data pada dataset tersebut adalah *items* dan *attributes*. *Dataset* tersebut termasuk dalam tipe data *items* karena entitas individunya bersifat diskrit dan ditampilkan sebagai *row* dalam tabel. Sementara itu, dataset tersebut juga termasuk dalam tipe data *attributes* karena terdapat beberapa *property* yang dapat diamati, diukur pada dataset tersebut yang ditampilkan sebagai kolom dalam tabel. Selain itu, setiap *cell* dalam tabel dispesifikasikan sebagai kombinasi antara *row* dan *column* (*item* dan *attributes*) dan di dalam *cell* tersebut terdapat nilai.

b. Tipe dataset (dataset type)

Tipe *dataset* Video Game Sales adalah *Tables*, karena pada *dataset* tersebut terdapat tabel sederhana dimana setiap baris merepresentasikan item data dan setiap kolomnya adalah atribut dari *dataset*.

c. Ketersediaan dataset (dataset availability)

Ketersediaan *dataset* (*dataset availability*) pada *dataset* Video Game Sales adalah *dynamic* karena diperoleh secara *online*, dimana ketika ingin mengakses *dataset* tersebut maka diperlukan adanya koneksi internet untuk dapat mengakses *dataset* tersebut dan mengunduhnya.

d. Tipe setiap atribut (attribut type)

Tabel 1 Tipe Atribut Pada Video Game Sales

Atribut	Attribut Type
<i>Rank</i>	<i>Ordered-quantitative</i>
<i>Name</i>	<i>Categorical</i>
<i>Platform</i>	<i>Categorical</i>
<i>Year</i>	<i>Quantitative</i>
<i>Genre</i>	<i>Categorical</i>
<i>Publisher</i>	<i>Categorical</i>
<i>NA_Sales</i>	<i>Ordered-quantitative</i>
<i>EU_Sales</i>	<i>Ordered-quantitative</i>
<i>JP_Sales</i>	<i>Ordered-quantitative</i>
<i>Other_Sales</i>	<i>Ordered-quantitative</i>
<i>Global_Sales</i>	<i>Ordered-quantitative</i>

e. Task

Task yang digunakan adalah membandingkan (*compare*) dimana visualisasi tersebut akan membandingkan setiap penjualan Video Game mana yang paling laris tiap tahunnya dalam waktu 5 tahun terakhir menurut negara Jepang. Selain *Compare task* yang digunakan adalah melihat *trend* Game Video apa saja yang terpopuler dalam waktu 10 tahun terakhir berdasarkan *Name Game* dan terakhir berdasarkan hasil produksi perusahaan (*Publisher*) Nintendo.

Untuk idiom visualisasi yang sesuai task *compare* adalah *bar chart* karena idiom visualisasi *bar chart* terdapat 2 atribut yang terlibat dalam visualisasi tersebut, yaitu Name *Game* sebagai *categorical* atribut dan Tahun (*year*) sebagai *quantitative*. Dimana yang berperan sebagai atribut *key* adalah Name *Game*, dan ketika *key* atributnya adalah *categorical* maka tipe visualisasi yang tepat untuk meng gambarkannya adalah *bar chart*.

Untuk idom visualisasi untuk task trend adalah *bar chart* karena Pada *bar chart* setiap data poin diwakili oleh batang dengan ketinggian tertentu sesuai dengan nilai data yang hendak diwakili. *Bar chart* dapat membandingkan nilai dari satu kategori dengan beberapa kategori lainnya dan melihat frekuensi atau besaran dari kategori/variabel tersebut. Misalnya untuk memperoleh daftar Video *Game* terpopuler menurut Name *Game* pada 10 tahun terakhir dari dataset Video *Game Sales*, maka Video *Game* yang menurut Name *Game* harus dibandingkan terlebih dahulu nilainya untuk melihat frekuensi nilai terbesar yang akan terlihat dari ketinggian bar nya dengan menggunakan idiom *bar chart* maka informasi pada *dataset* tersebut dapat dikomunikasikan dengan efektif dan efisien.

3. Evaluasi Visualisasi Data

Evaluasi Visualisasi Data perlu dilakukan, adapun evaluasi diperlukan dikarenakan :

- Tidak cukup untuk memunculkan ide baru
- Harus juga menentukan apakah metode baru lebih baik dari metode yang sudah ada
- Tidak ada evaluasi = tidak ada bukti kemajuan (Tidak tahu kapan kita berhasil)

	QUALITATIVE	QUANTITATIVE
FORMATIVE	Focus groups Market research	Usability studies
SUMMATIVE	Expert reviews Cognitive walkthroughs Think-aloud protocol Structured interviews Grounded theory Observational studies Insight-based evaluation	Controlled user studies Comparative user studies

Gambar 2 Tipe Evaluasi

Adapun evaluasi terdapat berbagai macam sesuai dengan kebutuhannya, hal tersebut dapat dilihat pada gambar diatas. Evaluasi kuantitatif merupakan evaluasi yang paling populer dimana : hasil dapat dengan mudah diukur, mudah dibandingkan dengan hasil

lain, sangat populer dalam penelitian HCI. Sedangkan untuk evaluasi kualitatif lebih sulit untuk dianalisis, sering tidak ada kesimpulan yang jelas, serta tidak ada angka. Berdasarkan idiom visualization yang terlkah ditentukan untuk task trend adalah *bar chart*. Kemudian untuk tipe evaluasi yang saya ajukan adalah Evaluasi Quantitative - Summative. Adapun karena hasil visualisasi tersebut agar mudah untuk dipahami, mudah dibanngan dengan data yang lain, juga populer dalam research HCI. Evaluasi Quantitative menggunakan sejumlah variabel yang melibatkan pengukuran yang dinyatakan dalam angka yang kemudian dianalisis untuk memastikan hasil akhir yang kemudian digeneralisasikan untuk memastikan bahwa teori yang digunakan adalah sesuai. Adapun variable attribute yang dianalisis yaitu dengan mnggundakan attribute yang telah disediakan pada dataset sesuai dengan idiom yang telah dibuat sebelumnya, setelah itu kemudian dilakukan analisis terhadap visualisasi apakah sudah sesuai dengan dataset kemudian terakhir dapat dilakukan pengambilan kesimpulan atas visualisasi tersebut.

3. Rencana Kerja

Pada bab Rencana Kerja akan dijelaskan mengenai jadwal kegiatan yang dibuat dengan menggunakan *Gantt Chart* dan Pembagian Tugas.

3.1 Jadwal Kegiatan

Berikut merupakan jadwal kegiatan dalam kegiatan pelaksanaan proyek, jadwal kegiatan pelaksanaan proyek dilampirkan dengan menggunakan *Gantt Chart*, yaitu:

Tabel 2 Jadwal Kegiatan Proyek

Daftar Kegiatan	Bulan ke -					
	November 2020		Desember 2020		Januari 2021	
	Week 11	Week 12	Week 13	Week 14	Week 15	Week 16
Menentukan Topik dan Mencari Dataset						
Mengerjakan Proposal						
Implementasi						
Mengerjakan Laporan Akhir						
Finalisasi Proyek						

3.2 Pembagian Tugas

Berikut merupakan pembagian tugas dalam kegiatan pelaksanaan proyek.

Tabel 3 Pembagian Tugas

Nama	Peran	Tugas
Rezky Simanjutak	Ketua Tim	Bertanggung jawab untuk mengkoordinasi pengerjaan proyek, mengimplementasi kode untuk membangun sistem.
Siti Berliana Manurung	<i>System Analyst</i>	Bertanggung jawab untuk memproses dan menganalisis data yang sudah dikumpulkan, serta melakukan pengujian terhadap sistem yang dibangun.
Mega Christy Silalahi	<i>System Analyst</i>	Bertanggung jawab untuk memproses dan menganalisis data yang sudah dikumpulkan, serta melakukan pengujian terhadap sistem yang dibangun.

4. Referensi

- [1] S.S.Stevens, On the Theory of Scales of Measurement, 1946.
- [2] B. Bowman, "Toward Visualization for Games: Theory, Design Space, and Patterns," *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, vol. XX, pp. 1-4, 2012.
- [3] T.Munzer, Visualization Analysis and Design,, CRC Press, 2015.