

BAB II

TINJAUAN PUSTAKA DAN DASAR TEORI

2.1 Tinjauan Pustaka

Beberapa jurnal atau penelitian yang terkait dengan penelitian ini adalah sebagai berikut :

Faristin (2005) membahas perancangan data warehouse yang dapat mendukung kebutuhan *end user* di dalam kegiatan akademik Universitas Indonesia dan juga dapat mendukung aplikasi data mining pada penelitian Fatma (2005). Selain itu penelitian tersebut juga bertujuan untuk mengetahui bagaimana teori langkah-langkah penerapan data warehouse dalam prakteknya. Pada penelitian tersebut data bersumber dari Sistem Informasi Akademik dan Kemahasiswaan Universitas Indonesia (SIAK UI).

Azimah (2007) membahas perancangan data warehouse dan penerapan teknik data mining untuk data akademik di Universitas Nasional untuk menggali informasi-informasi penting dan membangun model yang dapat membantu operasional sehari-hari agar dapat memberikan pelayanan yang terbaik buat mahasiswa. Hasil dari penelitian tersebut adalah pengembangan data warehouse dengan media presentasi aplikasi berbasis web. Sedangkan untuk analisa data menggunakan teknik data mining menghasilkan pola karakteristik mahasiswa yang mengambil suatu program peminatan tertentu.

Uyar dan Gungormus (2011) membahas faktor yang terkait dengan performa mahasiswa dalam mata kuliah akuntansi, delapan variabel telah

ditentukan yang kemungkinan memiliki dampak pada performa mahasiswa, antara lain nilai akuntansi, jenis kelamin, umur, nilai UN, nilai masuk universitas, jenis asal sekolah, nilai IPK, kehadiran dan nilai matematika. Hasil penelitian menunjukkan bahwa jenis kelamin dan nilai masuk universitas tidak ada hubungan korelasi dengan performa mahasiswa. Sedangkan pada usia memiliki pengaruh signifikan yang negatif. Sedangkan yang lainnya memiliki pengaruh terhadap performa mahasiswa. Pada penelitian tersebut data diambil dari sampel 114 mahasiswa manajemen dan 53 mahasiswa akuntansi di Istanbul Vocational School of Fatih University.

Radityo Adi Nugraha (2008) membahas analisa penjualan mobil bekas berbasis Multidimensional dan Star Schema Design. Data *warehouse* penjualan mobil ini, data data penjualan lebih terangkum ke dalam bentuk laporan yang detail berdasarkan dimensi yang disaring sebelumnya. Pada sistem data *warehouse* penjualan mobil ini juga dapat menampilkan laporan berupa grafik yang menggambarkan mengenai data *trend* penjualan mobil baik itu dari sisi *trend* warna, tipe mobil maupun *trend* penjualan tiap bulannya.

Eka Miranda (2011) membahas Implementasi Data Warehouse Dan Data Mining mengenai Analisis Peminatan Studi Siswa. Data Warehouse dapat digunakan untuk menyimpan data dalam jumlah besar dan juga berpotensi untuk mendapatkan sudut pandang baru dari sebaran data, dan memungkinkan untuk memberikan pelaporan dan jawaban dari pertanyaan pengguna yang bersifat ad hoc termasuk juga untuk melakukan analisis terhadap record akademik siswa untuk mendukung pengambilan keputusan pemilihan program

peminatan siswa. Proses transformasi dari database yang menyimpan data transaksi akademik siswa mampu memindahkan data penting yang berguna bagi universitas untuk melakukan analisis.

Implementasi data warehouse dan teknik data mining pada universitas dapat membantumengidentifikasi karakteristik pola siswa dalam mengambil program peminatan yang ada di jurusan Sistem Informasi. Hasilnya adalah diperoleh karakteristik pola siswa yang memilih program peminatan CIS, EBIS dan Applied Databases dilihat dari prestasi akademiknya, yang pada akhirnya dapat membantu siswa lain dalam mengambil keputusan.

Secara detail perbedaan penelitian ini dengan penelitian yang sudah ada dapat dilihat pada tabel berikut :

Tabel 2. 1 Perbandingan Penelitian

Penulis (Tahun)	Topik Penelitian	Keterangan	Tools
Faristin (2005)	Perancangan data warehouse untuk kegiatan akademik di Universitas Indonesia	Database : Oracle Analisis : Data Warehouse Domain : Data Akademik	- Transparent Gateway (Oracle) - Oracle Replication
Azimah (2007)	Perancangan data warehouse dan penerapan teknik data mining untuk data akademik	Database : PostgreSQL Analisis : Data Warehouse dan Penerapan teknik Data Mining Domain : Akademik	- Mondrian - JPivot
Uyar dan Gungormus (2011)	Faktor terkait yang berhubungan dengan performa mahasiswa	Analisis : Hubungan korelasi yang	- Microsoft Excel

	di matakuliah Financial Accounting	mempengaruhi performa mahasiswa Domain : Mata kuliah akuntansi	- SPSS
Radityo Adi Nugraha (2008)	Aplikasi Data Warehouse untuk Analisis Penjualan Mobil Berbasis Multidimensional Modeling(MDM) dan Star Schema Design (Studi Kasus PT.Asco Automotive)	Analisis: Data Warehouse dan penerapan berbasis Multidimensional Modeling(MDM) dan Star Schema Design	Visual Basic, .Net 2005 dan SQL Server 2000
Eka Miranda (2011)	Implementasi Data Warehouse Dan Data Mining: Studi Kasus Analisis Peminatan Studi Siswa	Analisis : Data Warehouse dan Penerapan teknik Data Mining Domain : Akademik	ETL Tool(Extraction, Transformation, Loading)
Atmaja (2019)	Perancangan data Warehouse beasiswa Bidikmisi	Database : Mysql Analisis : Data Warehouse Domain : Data Beasiswa	Microsoft Excel

Berbeda dengan penelitian sebelumnya, pada penelitian ini akan dilakukan perancangan dan pembuatan data warehouse UGM untuk beasiswa bidikmisi di UGM. Data yang digunakan bersumber dari Direktorat Kemahasiswaan UGM. Penelitian ini juga bertujuan untuk memperoleh informasi yang bermanfaat untuk level eksekutif seperti melihat jumlah mahasiswa penerima bidikmisi, rata-rata IPK, rata-rata lama studi berdasarkan fakultas, program studi dan asal daerah. *Tool* yang digunakan adalah Microsoft Excel. Kemudian dari hasil data warehouse yang telah dibangun akan dibuat aplikasi data mining yang berguna

untuk proses analisis korelasi yang berguna untuk melihat hubungan korelasi antara domain IPK lulus, lama studi, penghasilan orangtua berdasarkan fakultas, prodi, asal daerah dan angkatan.

2.2 Dasar Teori

Dasar teori digunakan untuk memahami definisi, pengertian dasar dan istilah yang digunakan dalam penelitian ini. Berikut dasar teori yang digunakan

2.2.1 Data Warehouse

Data warehouse adalah sistem *repository* informasi yang didapatkan dari berbagai sumber, disimpan dalam skema yang sama, pada satu sisi dan digunakan untuk mendukung proses pengambilan keputusan.

1. Model Data Warehouse

Data warehouse bersumber dari berbagai operasional database (*OnLine Transaction Processing / OLTP* database). Kemudian data yang ada dalam data warehouse dapat dimanfaatkan oleh aplikasi analisa dengan menggunakan *query*, sehingga menghasilkan suatu *strategic information*. Pada akhirnya informasi tersebut digunakan untuk mendukung proses analisa bisnis dan pengambilan keputusan.

2. Karakteristik Data Warehouse

King (2000) mengatakan bahwa data warehouse memiliki beberapa karakteristik, yaitu *subject-oriented*, *integrated*, *time variant*, *non-volatile* dan *data granularity*.

Pada *operational systems* data disimpan berdasarkan aplikasi. Set data hanya terdiri dari data yang dibutuhkan oleh fungsi yang terkait dengan aplikasinya. Sedangkan pada data warehouse data disimpan bukan berdasarkan aplikasinya melainkan berdasarkan subjek datanya. Subjek data yang dimaksud disini adalah subjek bisnis yang penting bagi perusahaan atau instansi terkait.

Misalnya untuk sebuah perusahaan manufaktur subjek bisnis yang penting, yaitu penjualan, pengangkutan dan penyimpanan barang. Menurut Connolly and Begg (2005), terdapat perbedaan antara sistem OLTP dan sistem data warehouse, seperti ditunjukkan pada Tabel 2.

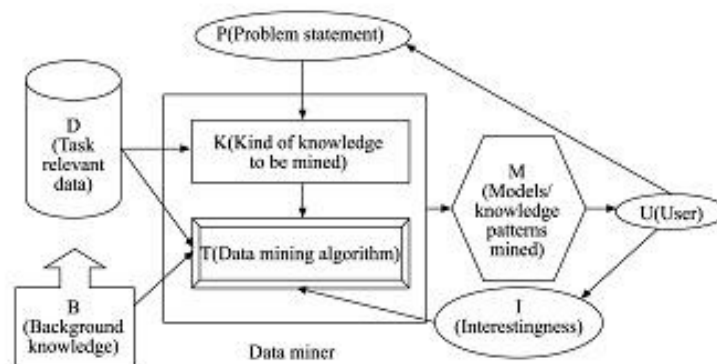
Tabel 2. 2 Perbandingan OLTP dan Data Warehouse

OLTP	Data warehouse
Menangani data sekarang	Menangani data historis
Data bersifat dinamis	Data bersifat statis
Proses berulang	Proses sewaktu-waktu tidak terstruktur dan <i>heuristic</i>
Jumlah transaksi tinggi	Jumlah transaksi rendah sampai sedang
<i>Transaction driven</i>	<i>Analysis driven</i>
Berorientasi aplikasi	Berorientasi subjek
Mendukung untuk keputusan harian	Mendukung untuk keputusan strategis
Banyak <i>user</i>	Sedikit <i>user</i> (tingkat manajer)

2.2.2 Data Mining

Data mining (DM) adalah proses mencari dan menggali hubungan antar data yang berguna bagi penggunaannya. Data mining (DM) juga merupakan proses

mencari data untuk menemukan yang sebelumnya tidak diketahui dari hubungan antara data yang menarik untuk penggunaannya (Han dan Kamber, 2001). Tahap terpenting dalam mendapatkan pengetahuan dari dalam database adalah data mining . DM telah menjadi bidang ilmu tersendiri (Liao dan Chu, 2012). Data mining adalah subjek yang terus berkembang (Liao dan Chu, 2012) di mana digunakan untuk mendapatkan pengetahuan (knowledge discovery) dari dalam database yang mengacu pada keseluruhan proses mulai dari low-level data sampai ke high-level data. Data mining adalah proses untuk menemukan suatu pola dan trend dari data. Tujuan dari proses ini adalah menyusun data yang berjumlah besar untuk mendapatkan informasi baru. Data mining juga dapat berarti eksplorasi dan analisis data secara otomatis atau semi otomatis dari data yang jumlahnya besar untuk mendapatkan pola dan aturan yang bermakna bagi penggunaannya (Fayyad, U., Shapiro, G.P. dan Smyth, P., 1996). menjelaskan bahwa proses data mining terdiri atas enam proses kerja utama yang dapat dilihat pada Gambar 1



Gambar 2. 1 Data Mining

Data mining adalah proses mencari trend dan pola dalam data. Tujuan dari proses ini adalah untuk menyortir besar jumlah data dan menemukan informasi

baru. Manfaat dari data mining adalah pengetahuan baru ditemukan dan hasil ditindaklanjuti. Tujuan dari data mining adalah untuk mengekstrak informasi yang berharga dan baru dari data yang ada. Secara umum, data mining meliputi fungsi utama sebagai berikut: klasifikasi, clustering, estimasi, prediksi, dan deskripsi. Teknologi data mining dapat dibagi antara teknologi tradisional dan halus. Statistik analisis merupakan perwakilan dari teknologi tradisional. Adapun teknologi data mining halus adalah semua jenis kecerdasan buatan yang digunakan, yang lebih sering digunakan adalah jenis pohon keputusan, jaringan syaraf, algoritma genetika, logika fuzzy dan aturan induksi. Penggunaan berbagai jenis aplikasi yang berbeda sering dapat menyebabkan hasil yang berbeda.

2.2.3 Korelasi Rho Spearman

Korelasi Spearman merupakan alat uji statistik yang digunakan untuk menguji hipotesis asosiatif dua variabel bila datanya berskala ordinal (ranking) (Sugiono, 2004). Nilai korelasi ini disimbolkan dengan ρ (dibaca: rho). Karena digunakan pada data berskala ordinal, untuk itu sebelum dilakukan pengolahan data, data kuantitatif yang akan dianalisis perlu disusun dalam bentuk ranking.

Nilai korelasi Spearman berada diantara $-1 < \rho < 1$. Bila nilai $\rho = 0$, berarti tidak ada korelasi atau tidak ada hubungannya antara variabel independen dan dependen. Nilai $\rho = +1$ berarti terdapat hubungan yang positif antara variabel independen dan dependen. Nilai $\rho = -1$ berarti terdapat hubungan yang negatif antara variabel independen dan dependen. Dengan kata lain, tanda “+” dan “-” menunjukkan arah hubungan di antara variabel yang sedang dioperasikan.

Uji signifikansi Spearman menggunakan uji Z karena distribusinya mendekati distribusi normal. Kekuatan hubungan antara variabel ditunjukkan melalui nilai korelasi. Koefisien korelasi diinterpretasikan (De Vaus:2002) sebagai berikut:

Tabel 2. 3 Makna Korelasi Rho Spearman

NILAI	MAKNA
0,00 - 0,19	Sangat rendah/sangat lemah
0,20 - 0,39	Rendah/lemah
0,40 - 0,59	Sedang
0,60 - 0,79	Tinggi/kuat
0,80 - 1,00	Sangat tinggi/kuat

Rumusan Korelasi Spearman adalah sebagai berikut:

$$\rho = 1 - (6\sum b) / (n(n^2 - 1))$$

Keterangan :

ρ : nilai korelasi rank spearman

b : jumlah kuadrat selisih ranking variabel x dan y atau $R_X - R_Y$

n : jumlah sampel

2.3 Tools

2.3.1 Microsoft Excel

Microsoft Excel adalah aplikasi spreadsheet canggih yang bisa digunakan untuk menampilkan data, melakukan pengolahan data, kalkulasi, membuat

diagram, laporan, dan semua hal yang berkaitan dengan data yang berupa angka (Wiki, 2018). Contoh aplikatif dari penggunaan spreadsheet dalam kehidupan sehari-hari misalnya untuk keperluan sebagai berikut :

1. Menghitung rata-rata atau nilai maksimum suatu data.
2. Membuat sebuah grafik yang memperlihatkan presentasi suatu penjualan dalam range tertentu.
3. Memperlihatkan jumlah total suatu variabel.
4. Memanajemen suatu database.

Spreadsheet pada dasarnya adalah grid besar yang menata data ke dalam baris dan kolom, namun Excel memiliki fitur yang lebih dan hanya sekedar pengganti buku akuntansi. Excel dapat melakukan semua kalkulasi yang diinginkan, dan misalkan data di dalam sebuah sheet Excel diubah/di-update, maka Excel dapat langsung melakukan peng-update-an hasil tanpa harus mengubah struktur keseluruhan dari penulisan spreadsheet.

Salah satu software spreadsheet saat ini adalah Microsoft Excel. Aplikasi pengolah angka ini merupakan salah satu bagian dalam Microsoft Office. Menurut penelitian mengenai penggunaan program di Amerika dinyatakan bahwa Microsoft Excel keluaran dari Microsoft Corporation ini terbukti paling laris dan paling banyak digunakan di pasaran jika dibandingkan dengan program pengolah angka sejenis seperti Quattro, Quattro Pro, Lotus, dan Supercall. Hal tersebut dikarenakan Microsoft Excel berada di bawah sistem operasi Windows yang memiliki keunggulan berbasis grafik (Grafic user interface/GUI) yang dikembangkan oleh Microsoft, artinya sistem ini selain digunakan untuk mode

grafis juga memiliki kelebihan untuk melaksanakan tugas ganda/multitasking. Pada saat yang bersamaan dapat digunakan untuk menjalankan program lain yang berbeda.

Microsoft Office Excel merupakan pengembangan dan Microsoft Office versi sebelumnya dan merupakan salah satu produk Microsoft Corporation. Pada Excel tersedia berbagai macam kemampuan untuk membantu menyelesaikan berbagai pekerjaan, khususnya yang berhubungan dengan pengolahan angka, baik dalam perkantoran maupun dalam dunia bisnis lainnya.

2.3.2 Hypertext Markup Language (HTML)

Hypertext Markup Language (HTML) adalah bahasa yang digunakan untuk menulis halaman web. HTML merupakan pengembangan dari standar pemformatan dokumen teks yaitu Standard Generalized Markup Language (SGML) (Kadir, 2013). HTML sebenarnya adalah dokumen ASCII atau teks biasa, yang dirancang untuk tidak tergantung pada suatu sistem operasi tertentu.

Mendesain HTML berarti melakukan suatu tindakan pemrograman. Namun HTML bukanlah sebuah bahasa pemrograman. Namun HTML hanyalah berisi perintah-perintah yang telah terstruktur berupa tag-tag penyusun. Menuliskan tag-tag HTML tidaklah sebatas hanya memasukkan perintah-perintah tertentu agar HTML kita dapat di akses oleh browser. Mendesain HTML adalah adalah sebuah seni tersendiri. Homepage yang merupakan implementasi dari HTML adalah refleksi dari orang yang membuatnya. Untuk itu kita perlu mendesainnya dengan baik agar para pengunjung homepage yang kita buat

merasa senang dan bermanfaat. Mendesain HTML dapat dilakukan dengan cara menggunakan HTML Editor, seperti notepad++, adobe dreamweaver dan lain-lain.

Dengan cara menuliskan sendiri secara manual satu persatu tag-tag HTML ke dalam dokumen HTML. Ada kelebihan dan kekurangan dari dua cara di atas. Cara pertama kelebihanannya adalah HTML Editor merupakan sebuah program yang khusus didesain untuk membuat, melakukan editing bahkan mem-publish ke internet. Dengan kemampuannya menggabungkan kemudahan dan kecanggihan teknologi internet ke dalam dokumen HTML maka cara ini sangat disukai oleh para pemula dan desainer yang tidak ingin belajar lebih mendalam mengenai HTML. Sedangkan cara kedua adalah menuliskan secara manual satu persatu tag-tag HTML. Hal ini sangat disarankan sulit dikarenakan akan memakan tenaga dan waktu ekstra untuk melakukannya, ditambah lagi Anda harus melakukan cara-cara konvensional untuk melihat hasilnya pada web browser. Namun pada cara kedua adalah dasar dari segala bentuk HTML yang akan Anda pelajari, karena dengan cara itulah Anda akan lebih paham mengenai cara kerja dan berbagai perintah yang biasa dipakai pada bahasa HTML.

2.3.3 PHP

PHP (*HyperText Preprocessor*) adalah sebuah bahasa utama *script server side* yang disisipkan pada HTML yang dijalankan di *server*. Menurut Betha Sidik, dalam bukunya yang berjudul Pemrograman Web Dengan PHP (2012 : 4), menyebutkan bahwa : “PHP merupakan secara umum dikenal dengan sebagai

bahasa pemrograman *script-script* yang membuat dokumen HTML secara *on the fly* yang dieksekusi di *server* web, dokumen HTML yang dihasilkan dari suatu aplikasi bukan dokumen HTML yang dibuat dengan menggunakan editor teks atau editor HTML, dikenal juga sebagai bahasa pemrograman *server side*".

2.3.4 CSS

Cascading Style Sheet (CSS) merupakan aturan untuk mengendalikan beberapa komponen dalam sebuah web sehingga akan lebih terstruktur dan seragam. CSS bukan merupakan bahasa pemrograman. Sama halnya dengan *style* dalam aplikasi pengolahan kata misalnya Microsoft Word yang mengatur beberapa *style* seperti *heading*, *subbab*, *bodytext*, *footer*, *image* dan *style* lainnya untuk dapat digunakan bersama-sama dalam beberapa *file*.

2.3.5 Java Script

JavaScript adalah salah satu bagian dari bahasa pemrograman Java yang dibuat oleh Sun Microsystems. Javascript adalah bahasa script yang langsung dimasukkan ke dalam dokumen HTML, sehingga tidak memerlukan kompiler lagi (Kadir, 2013)

2.3.6 MySQL

Menurut Betha Sidik, dalam bukunya yang berjudul Pemrograman Web Dengan PHP (2012 : 333), menyebutkan bahwa :“MySQL merupakan *software* database yang termasuk paling populer di lingkungan Linux, kepopuleran ini

karena ditunjang performasi *query* dari databasenya yang saat itu bisa dikatakan paling cepat dan jarang bermasalah”.