

Implementasi Extract Transform Load Pada Data Akademik Perguruan Tinggi

Nahrin Hartono¹, Didit Hermawan²

^{1,2}Jurusan Sistem Informasi, Universitas Islam Negeri Alauddin Makassar, Indonesia
nahrinhartono@gmail.com¹, didithrwn09@gmail.com²

Informasi Artikel

Article historys:

Dipublikasi Jun 30, 2021

Kata Kunci:

Akademik
 Pentaho
 Extract
 Transform
 Load

ABSTRACT

Each department at university needs to analyze student data and scores. By analyzing student grades, majors can improve lecturer performance in learning; one way to explore student grades is to collect student grade data and then process the data to produce information in the form of learning performance results that can be used to make decisions to improve lecturer performance. The author will analyze a structured programming course student score data; the following collected data are related to student attendance, assignment scores, mid-semester exam scores, and final semester exam scores. After all the data needed has been organized to build the data warehouse, the following process is the ETL process (Extract, Transform, and Load), as the tool used in the ETL process is Pentaho. This data extraction process using 1 data source, namely student value data with excel file type. After carrying out the extraction process, there is a data transformation process, namely making some changes to the extracted data to make it more consistent and uniform according to the needs of the data warehouse.

*Koresponden Author:

Nahrin Hartono,
 Jurusan Sistem Informasi,
 Universitas Islam Negeri Alauddin Makassar,
 Jl. H.M. Yasin Limpo No. 36 Samata, Kab Gowa, Sulawesi Selatan, Indonesia.

1. PENDAHULUAN

Penggunaan teknologi di bidang pendidikan untuk mengelola data nilai mahasiswa sudah sejak lama digunakan. Masalahnya, data yang dihasilkan tersebut tersebar di berbagai dosen sesuai dengan matakuliah yang ada sehingga menyulitkan operator jurusan ketika akan mengolah maupun menganalisis datanya. Untuk mengatasi masalah tersebut diperlukan data warehouse atau gudang data. Seperti namanya, Menurut Poe (1996:6), data warehouse adalah database analisis yang digunakan sebagai dasar sistem pengambilan keputusan. Analisis yang digunakan bersifat hanya untuk dibaca saja (*read only*) sehingga tidak dapat dilakukan perubahan data yang digunakan [1]. Pengertian data warehouse menurut Inmon (2002:31), Data warehouse adalah koleksi data yang mempunyai sifat berorientasi subjek, terintegrasi, tidak mengalami perubahan, dan mempunyai variasi waktu yang digunakan untuk mendukung proses pengambilan keputusan manajemen [2]. Jadi data warehouse diibaratkan gudang tempat penyimpanan informasi dan data nilai mahasiswa yang dikumpulkan dari berbagai dosen dan disimpan dalam skema terstruktur. Dengan demikian, keberadaan data warehouse sangat membantu dalam menganalisis kinerja pembelajaran dosen dan mahasiswa.

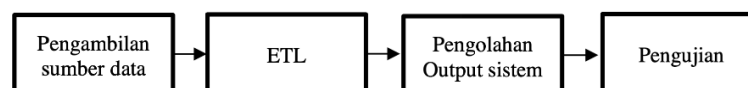
Pentaho adalah sebuah software yang digunakan dalam proses ETL dalam penyusunan data warehouse. Elemen utama dari Pentaho Data Integration (PDI) adalah Transformation dan Job.

Transformation adalah sekumpulan instruksi untuk merubah input menjadi output yang diinginkan (input-proses-output). Sedangkan Job adalah kumpulan instruksi untuk menjalankan transformasi [3]. Pentaho Community Edition atau Pentaho BI Suite adalah produk dari Pentaho yang sudah terintegrasi dengan kemampuan reporting, dashboard, data mining, workflow, dan ETL [4]. ETL adalah sekumpulan proses untuk mengambil dan memproses data dari satu atau banyak sumber data menjadi sumber baru [5]. ETL adalah proses inti dari integrasi data dan biasanya terkait dengan data warehouse. Tools ETL mengekstrak data dari sumber yang dipilih, mengubahnya menjadi format baru sesuai dengan aturan bisnis, dan kemudian memuatnya kedalam struktur data target [8]. ETL merupakan proses yang sangat penting dalam data warehouse, dengan ETL inilah data dari operational dapat dimasukkan kedalam data warehouse [6].

Pengajaran, penelitian dan pengabdian kepada masyarakat merupakan kinerja pokok dosen. Proses perkuliahan merupakan salah satu kinerja yang termasuk dalam kategori pengajaran. Proses perkuliahan pada perguruan tinggi di laksanakan selama satu semester umumnya berlangsung selama enam bulan. Masing masing perguruan tinggi telah menetapkan aturan penilaian tersendiri. Penilaian mahasiswa dilakukan oleh masing masing dosen pengampuh mata kuliah dan merupakan hak penuh dosen untuk memberikan nilai kepada mahasiswa berdasarkan proses pembelajaran yang telah dilalui. Untuk melakukan evaluasi terhadap kinerja dosen, salah satu yang dilakukan oleh pihak jurusan adalah dengan mengevaluasi proses perkuliahan yang telah dilaksanakan. Evaluasi dilakukan dengan melihat dan menganalisis dokumen dokumen perkuliahan seperti absesn perkuliahan dan monitoring perkuliahan serta penilaian akhir mahasiswa. Nilai akhir mahasiswa dituliskan berupa huruf yang umunya disebut nilai mutu, nilai akhir sangat berpengaruh terhadap lama studi mahasiswa sehingga jurusan perlu untuk melakukan pemetaan nilai mutu mahasiswa, hal ini dilakukan sebagai data awal untuk memprediksi rata rata lama studi mahasiswa.

2. METODE PENELITIAN/ALGORIMA

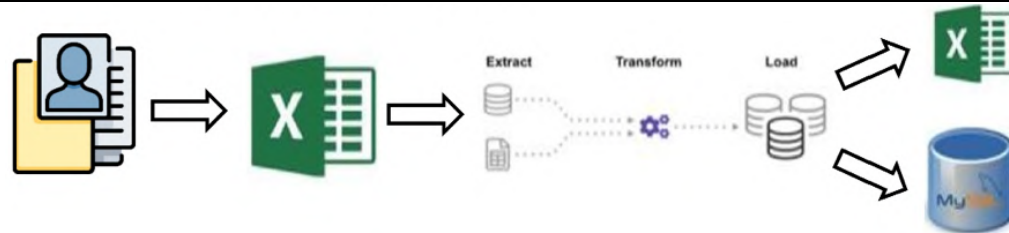
Penelitian ini merupakan case studies research yang merupakan penelitian yang memusatkan perhatian pada suatu kasus tertentu dengan menggunakan individu atau kelompok sebagai bahan penelitian. Objek penelitian adalah data nilai mahasiswa, adapun data tersebut diambil dari nilai pada satu mata kuliah yang diajarkan pada salah satu jurusan pada perguruan tinggi negeri di Makassar dengan jumlah mahasiswa perkelas rata rata 30 orang. Adapun tahapan penelitian yang dilakukan dapat dilihat pada gambar 1 berikut:



Gambar 1. Tahapan penelitian

Gambar 1 menunjukkan tahapan penelitian yang dilakukan, yang pertama adalah tahap pengambilan sumber data, sumber data pada penelitian ini adalah data nilai mahasiswa. Selanjutnya data nilai dilakukan Extract, Transform dan Load (ETL) yang merupakan tahap utama pada penelitian ini. Proses ETL dilakukan dengan menggunakan aplikasi Pentaho Data Integration (PDI). PDI merupakan aplikasi tersedia dengan versi berbayar dan versi komunitas yang bisa digunakan secara gratis. Pada tahap pengolahan output sistem akan didesain skema ETL dengan yang nantinya menghasilkan dua luaran. Tahap selanjutnya adalah pengolah output dan pengujian, pengolahan output sistem di lakukan dengan menyimpan data luaran ke file dengan format excel dan juga disimpan kedalam sistem database. Database yang digunakan pada penelitian ini menggunakan MySQL. Proses pengujian nantinya akan dilakukan dengan cara mengambil beberapa sampel data secara acak dan mencocokkan hasil antara pemrosesan data secara manual dengan hasil pemrosesan data melalui proses ETL di Pentaho Data Integration.

Secara umum penelitian ini menggambarkan proses ETL menjadi luaran file yang mudah dimengerti oleh pengguna, adapun gambaran umum proses penelitian yang dilakukan ditunjukkan pada gambar 2 berikut:



Gambar 2. Gambaran Umum Penelitian

3. HASIL PENELITIAN DAN PEMBAHASAN

Tahap pertama penelitian ini adalah pengolahan sumber data, seperti yang dijelaskan sebelumnya data yang digunakan pada penelitian ini adalah data nilai. Data nilai pada penelitian ini menggunakan format file excel seperti yang ditunjukkan pada gambar 3 berikut:

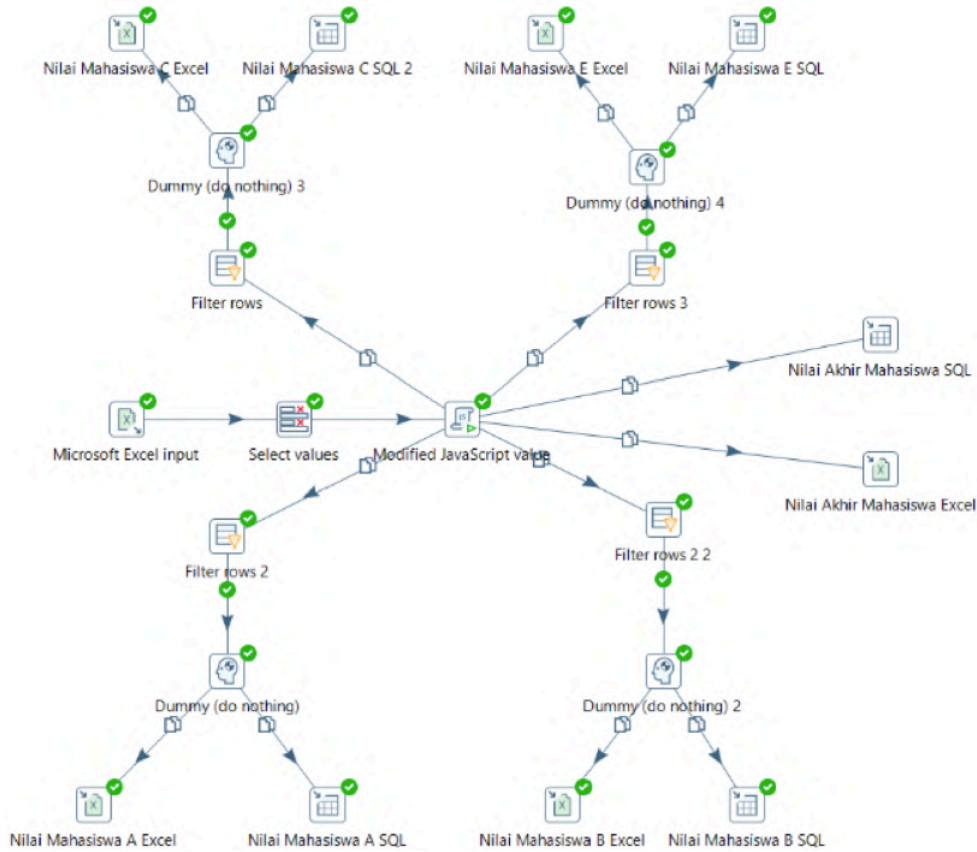
NIM	Nama	Kehadiran	Laporan	Tugas	MID	UAS
60900118032	MUH. HABIB MAJULANA	0	0	0	0	0
60900118032	ANDI NURUL FADHILLAH	100	88.3	91	87	90
60900118033	AHMAD ILHAM	90	75.8	82.5	78	74
60900118034	RESKI WAHIDANAH	100	80.8	79.5	78	75
60900118035	MAYANG SARI	100	80	80	80	78
60900118036	M. NURSYAHBANI	90	79.7	77.5	78	72
60900118037	WALYU HIBRAT	100	84.5	85	75	82
60900118038	FUJRI SUKI NAMDANI	100	80.3	83.3	85	76
60900118039	ILHAM PRABUJAYA	100	78.8	78	78	75
60900118040	AJMIANTI	100	84.2	85	80	78
60900118041	RIZKY SATRIA	90	81.8	84.3	80	76
60900118042	IRI RAHAYU PINI PURNAWATI	100	88	90.5	87	87
60900118043	REZA ARDIANSYAH	100	75.3	81.5	85	75
60900118044	DHIT HERMAWAN	100	88.3	88	83	88
60900118045	YUSRI	40	69.3	43.8	70	60
60900118046	ANDI ANISA	100	78.8	81.5	80	72
60900118047	IHYA ULUMUDDIN	100	82	83.8	85	75
60900118048	RIZKY AMALIA	80	75.8	85.5	80	80
60900118049	NAJAH MUHAMMAD	100	78	82	80	74
60900118050	DEVI LESTARI	100	82.3	83.3	80	78
60900118051	HADIARNI	100	81	80	80	70
60900118052	NURI ISMI	90	80.3	84.3	80	70
60900118053	FARRHAN	80	80.7	82	75	75
60900118054	MUSANTIS	100	81.2	85.5	85	77
60900118055	MAGFIRA	100	85.7	89	87	78

Gambar 3. Format Data Nilai

Gambar 3 menunjukkan format data nilai, dapat dilihat terdapat beberapa kolom yaitu nim, nama, kehadiran, laporan, tugas, mid dan uas. Untuk penilaian akhir mahasiswa didapatkan dari rata rata komponen penilaian. Dari data nilai tersebut kemudian akan dihasilkan laporan penilaian berikut:

1. Laporan nilai akhir mahasiswa
2. Laporan nilai mutu mahasiswa (A)
3. Laporan nilai mutu mahasiswa (B)
4. Laporan nilai mutu mahasiswa (C)
5. Laporan nilai mutu mahasiswa (E)

Proses ETL (Extract, Transform, Load) dilakukan menggunakan Tools Pentaho yang ditunjukkan pada gambar 4 berikut:



Gambar 4. Skema Proses ETL

Gambar 4 menunjukkan skema proses ETL, skema proses disusun sedemikian rupa dengan menggunakan beberapa step. Setiap step akan melakukan proses data yang akhirnya akan menghasilkan luaran yang diharapkan.

Transformasi data nilai akhir merupakan tahap untuk mendapatkan nilai akhir mahasiswa, untuk mendapatkan nilai akhir terlebih dahulu data nilai diinput ke dalam pentaho kemudian akan di select values data mana saja yang akan di ambil kemudian akan di modifikasi dengan Modified JavaScript Value, dan pada bagian akhir data akan di simpan dengan file Excel dan MySQL seperti ditunjukkan pada gambar 5 berikut:

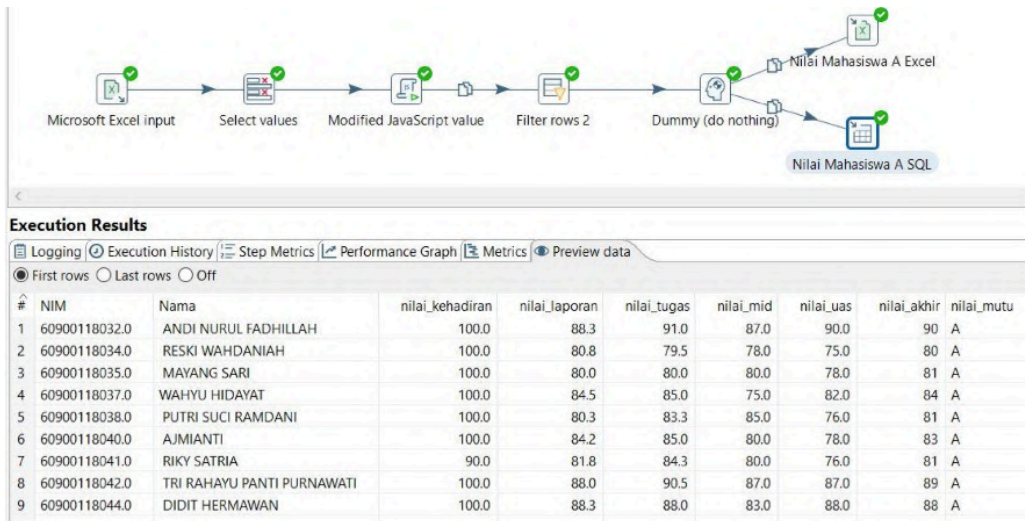
The screenshot shows a data transformation step in Pentaho Data Integration. The process flow is: Microsoft Excel input -> Select values -> Modified JavaScript value. The output of 'Modified JavaScript value' is split into two destinations: 'Nilai Akhir Mahasiswa SQL' and 'Nilai Akhir Mahasiswa Excel'.

Execution Results

#	NIM	Nama	nilai_kehadiran	nilai_laporan	nilai_tugas	nilai_mid	nilai_uas	nilai_akhir	nilai_mutu
1	60900118031.0	MUH. HABIB MAULANA	0.0	0.0	0.0	0.0	0.0	0	E
2	60900118032.0	ANDI NURUL FADHILLAH	100.0	88.3	91.0	87.0	90.0	90	A
3	60900118033.0	AHMAD ILHAM	90.0	75.8	82.5	78.0	74.0	77	B
4	60900118034.0	RESKI WAHDANIAH	100.0	80.8	79.5	78.0	75.0	80	A
5	60900118035.0	MAYANG SARI	100.0	80.0	80.0	80.0	78.0	81	A
6	60900118036.0	M. NURSYAHBANI	90.0	79.7	77.5	78.0	72.0	78	B
7	60900118037.0	WAHYU HIDAYAT	100.0	84.5	85.0	75.0	82.0	84	A
8	60900118038.0	PUTRI SUCI RAMDANI	100.0	80.3	83.3	85.0	76.0	81	A
9	60900118039.0	ILHAM PRABUJAYA	100.0	78.8	78.0	78.0	75.0	79	B

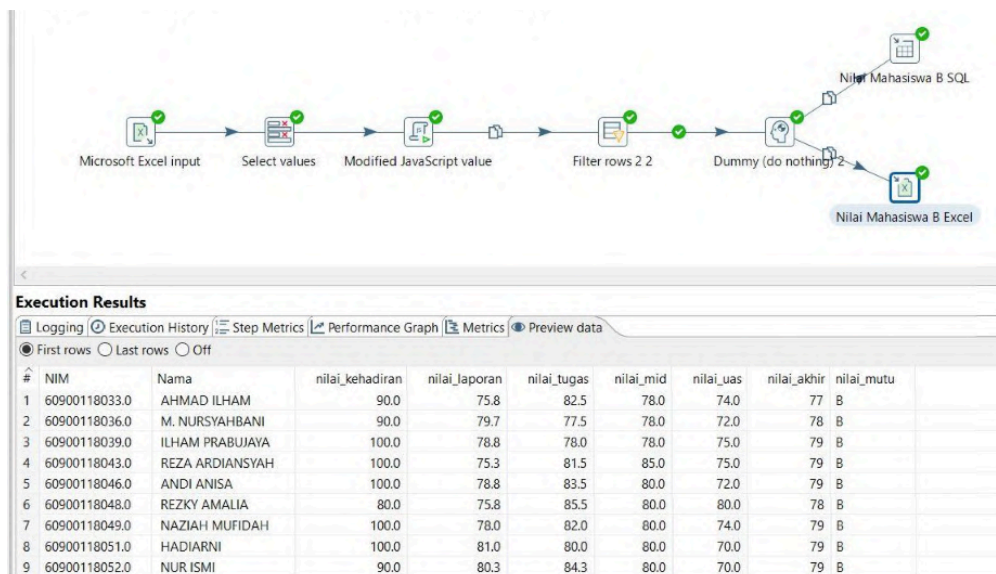
Gambar 5. Transformasi Data Nilai Akhir

Sedangkan untuk transformasi data nilai mutu A Modified JavaScript Value dimasukkan perintah untuk menghitung nilai mutu, selanjutnya akan di sortir berdasarkan nilai mutu A yang akan disimpan menjadi file Excel dan MySQL seperti ditunjukkan pada gambar 6 berikut:



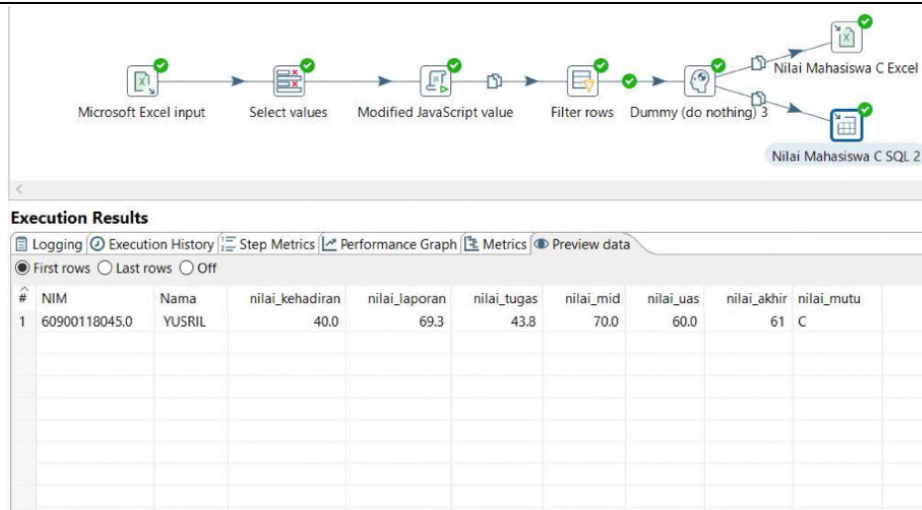
Gambar 6. Transformasi Data Nilai Mutu A

Transformasi nilai mutu B dilakukan dengan menginput data nilai terlebih dahulu ke dalam pentaho kemudian akan di select values data mana saja yang akan di ambil kemudian akan di modifikasi dengan Modified JavaScript Value, selanjutnya akan di sortir berdasarkan nilai mutu B yang akan disimpan menjadi file Excel dan MySQL seperti ditunjukkan pada gambar 7 berikut:



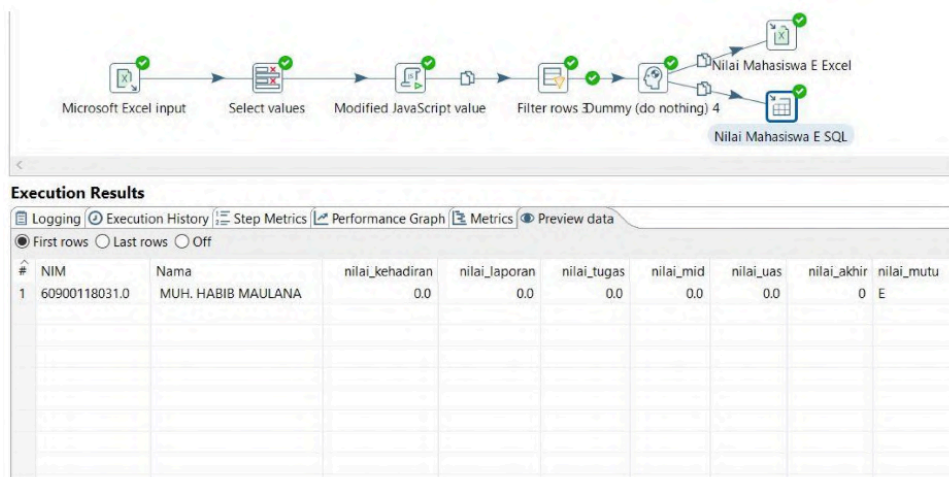
Gambar 7. Transformasi Data Nilai Mutu B

Transformasi nilai mutu C dilakukan dengan dengan menginput data nilai terlebih dahulu ke dalam pentaho kemudian akan di select values data mana saja yang akan di ambil kemudian akan di modifikasi dengan Modified JavaScript Value, selanjutnya akan di sortir berdasarkan nilai mutu C yang akan disimpan menjadi file Excel dan MySQL seperti ditunjukkan pada gambar 8. Tujuan Transform data ini agar dapat menjumlah hasil nilai akhir dan nilai mutu



Gambar 8. Transformasi Data Nilai Mutu C

Transformasi Data Nilai Mutu E dilakukan dengan menginput data nilai terlebih dahulu ke dalam pentaho kemudian akan di select values data mana saja yang akan di ambil kemudian akan di modifikasi dengan Modified JavaScript Value, selanjutnya akan di sortir berdasarkan nilai mutu E yang akan disimpan menjadi file Excel dan MySQL seperti ditunjukkan pada gambar 9 berikut:



Gambar 9. Transformasi Data Nilai Mutu E

Secara keseluruhan hasil dari pengolahan data nilai mahasiswa menggunakan pentaho adalah dengan dihasilkannya dua buah file output berupa file Microsoft Excel dan Database MySQL yang berisi data laporan nilai akhir, laporan nilai mutu A, laporan nilai mutu B, laporan nilai mutu C, dan Laporan nilai Mutu E. Laporan nilai akhir Mahasiswa merupakan rangkuman data nilai mahasiswa. Laporan ini akan menyajikan data nilai akhir dan nilai mutu yang merupakan penjumlahan dari data nilai kehadiran, nilai laporan, nilai tugas, nilai mid dan nilai uas seperti ditunjukkan pada gambar 10 berikut:

NIM	Nama	nilai_kahvejian	nilai_laporan	nilai_tugas	nilai_mid	nilai_uas	nilai_akhir	nilai_rmt
0900118031	MUH. HABIB MAULANA	0	0	0	0	0	0	E
0900118032	ANDI NURUL FADHILLAH	100	88.3	91	87	90	90	A
0900118033	AHMAD ILHAM	90	75.8	82.5	78	74	77	B
0900118034	RESKI WAHDANIAH	100	86.8	79.5	70	75	68	A
0900118035	MAYANG SARI	100	89	80	80	78	81	A
0900118036	M. NURSYAHBANI	90	79.7	77.5	70	72	78	B
0900118037	WAHYU HIDAYAT	100	84.5	85	75	82	84	A
0900118038	PURIS SUCI RAMDANI	100	80.3	83.3	85	76	81	A
0900118039	ILHAM PRABUJAYA	100	78.8	78	78	75	79	B
0900118040	ALMIANIT	100	84.2	85	80	78	83	A
0900118041	RIKY SATRIA	90	81.8	84.3	80	75	81	A
0900118042	TRI RAHAYU PANTI PURNAWATI	100	88	90.5	87	87	89	A
0900118043	REZA ARDIANSYAH	100	75.3	81.5	85	75	79	B
0900118044	DIDIT HERMAWAN	100	88.3	88	83	88	88	A
0900118045	YUSRIL	40	69.3	43.8	70	68	61	C
0900118046	ANDI ANISA	100	78.9	83.5	80	72	79	B
0900118047	BHYA ULUMUDDIN	100	82	83.8	85	75	82	A
0900118048	REZKY AMALIA	80	75.8	85.5	80	80	78	B
0900118049	NAZIAH MUFIDAH	100	78	82	80	74	79	B
0900118050	DEWI LESTARI	100	82.3	83.3	80	78	82	A
0900118051	HADIARNI	100	81	80	80	79	76	D
0900118052	NUR ISMI	90	80.3	84.3	80	79	78	B
0900118053	FAHRIANA	80	80.2	82	75	75	78	B
0900118054	MURANTI S	100	81.2	85.5	85	77	82	A
0900118055	MAGFIRA	100	85.7	89	87	78	85	A

Gambar 10. Laporan Nilai Akhir Mahasiswa (Excel dan MySQL)

4. KESIMPULAN

Implementasi metode ETL (Extract, Transform, Load) pada data bertujuan untuk menganalisis nilai mahasiswa. Analisis pada data ini nantinya akan digunakan sebagai laporan peningkatan kinerja pembelajaran bagi mahasiswa dan dosen. Pentaho Data Integration merupakan tools yang berguna untuk melakukan analisis data. Penulis sadar masih banyak kekurangan pada penelitian ini diantaranya adalah jumlah dan varian data yang digunakan masih sangat kurang. Untuk penelitian selanjutnya dapat melakukan dengan menambah jumlah data dan varian data.

DAFTAR PUSTAKA

- [1] Poe, Vidette. 1996. Building a Data Warehousing for Decision Support. New Jersey: Prentice Hall PTR.
- [2] Inmon, William H. 2002. Building the Data Warehouse. Third Edition. Canada: John Wiley and Sons.
- [3] Y. Wibisono, Pengantar Pentaho Data Integration (Kettle), vol. 5. 2014.
- [4] F. Zulfikar, Modul Pelatihan Fundamental Of Business Intelligence With Pentaho Community Edition. 2013.
- [5] JRP. Mulyana, Pentaho: Solusi Open Source untuk Membangun Data Warehouse. Bandung: Penerbit Andi, 2014.
- [6] V. Gour, S. S. Sarangdevot, G. S. Tanwar, and A. Sharma, "Improve Performance of Extract, Transform and Load (ETL) in Data Warehouse," Int. J. Comput. Sci. Eng., vol. 1, no. 3, pp. 786-789, 2010.