

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Data Cleansing pada data sistem rumah sakit membutuhkan beberapa metode untuk mengubah data kotor menjadi data yang berkualitas. Ditemukan bahwa data sistem rumah sakit mengalami duplikasi data pada tabel RAWAT_JALAN berjumlah 17 baris dan tabel PENDUDUK 6 baris, tidak konsistennya data pada tabel ASURANSI_KEPESERTAAN_VISIT berjumlah 55 baris, tabel KUNJUNGAN BPJS 3 baris, tabel PENDUDUK 381 baris, dan data kosong pada tabel KUNJUNGAN 137 baris dan RAWAT_JALAN 4 baris. Maka data sistem rumah sakit ini menerapkan metode *Duplicate Elimination*, *Inconsistency Detection*, dan *Handling Missing Entries*. Ketiga metode ini dapat dijalankan secara paralel atau berurutan. Dapat dijalankan untuk metode *Duplicate Elimination* terlebih dahulu lalu *Inconsistency Detection* dan setelahnya menggunakan *Missing Entries*.

Menggunakan metode *Duplicate Elimination* didukung oleh penelitian dari (Tamilselvi & Gifta, 2011) bahwa masalah mendeteksi dan menghilangkan data duplikat adalah salah satu masalah utama dalam pembersihan data dan kualitas data pada database. Penghapusan duplikat sulit karena disebabkan oleh beberapa jenis kesalahan seperti kesalahan tipografi, dan representasi berbeda dari nilai logis yang sama. Maka dari itu, dengan metode – metode tersebut data sistem rumah sakit sudah tidak mengalami duplikasi data, tidak konsistennya data dan tidak ada lagi data kosong.

5.2 Saran

- a. Selain menggunakan metode *Duplicate Elimination*, *Inconsistency Detection*, dan *Handling Missing Entries*, pembaca diharapkan dapat menghapus atribut yang memiliki data kosong seperti yang ditemukan pada penelitian ini
- b. Bagi pengembang sistem, dengan adanya penelitian ini membuktikan bahwa mengenai *data cleansing* pada data sistem rumah sakit merupakan hal yang penting untuk dilakukan sebelum proses *data mining* dilanjutkan agar menghasilkan data yang lebih akurat.