**Laporan Tugas PSD E**

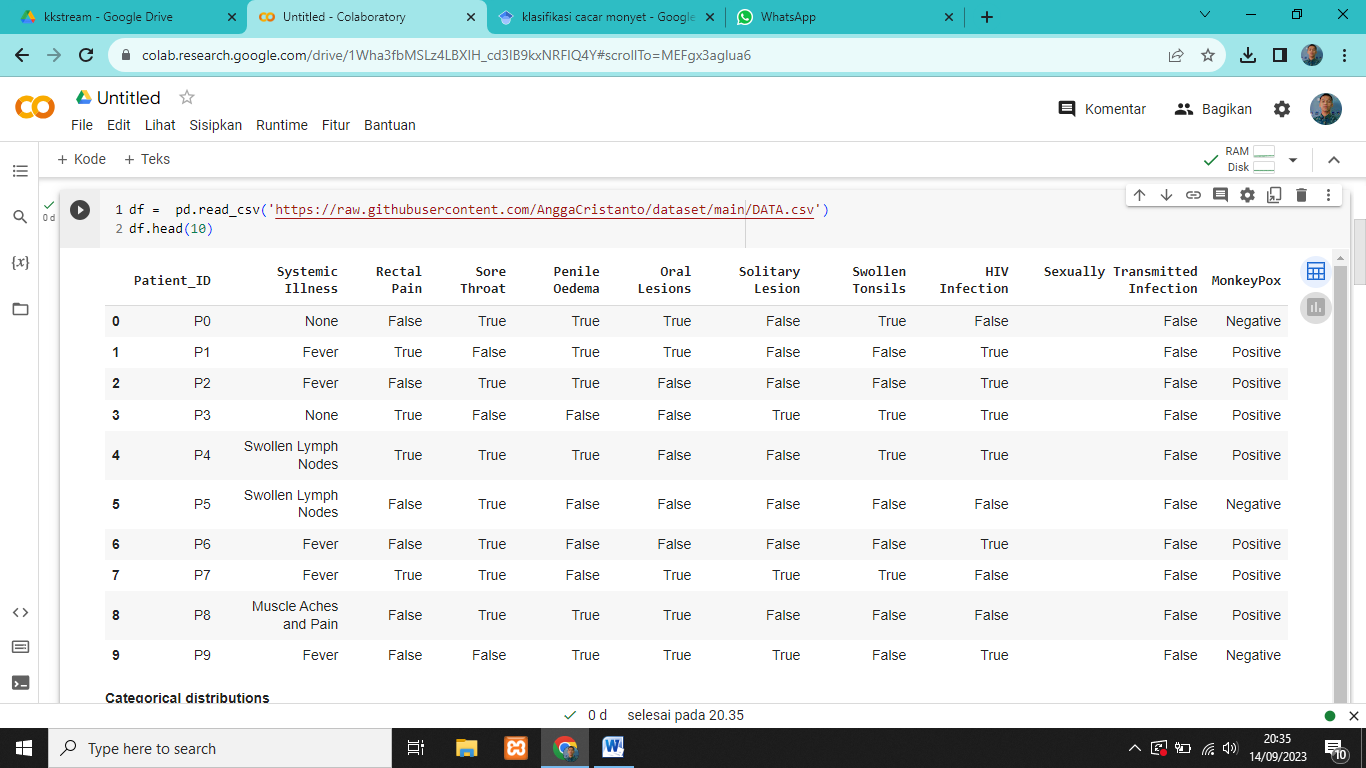
Nama : Angga Cristanto

Nim : 210411100039

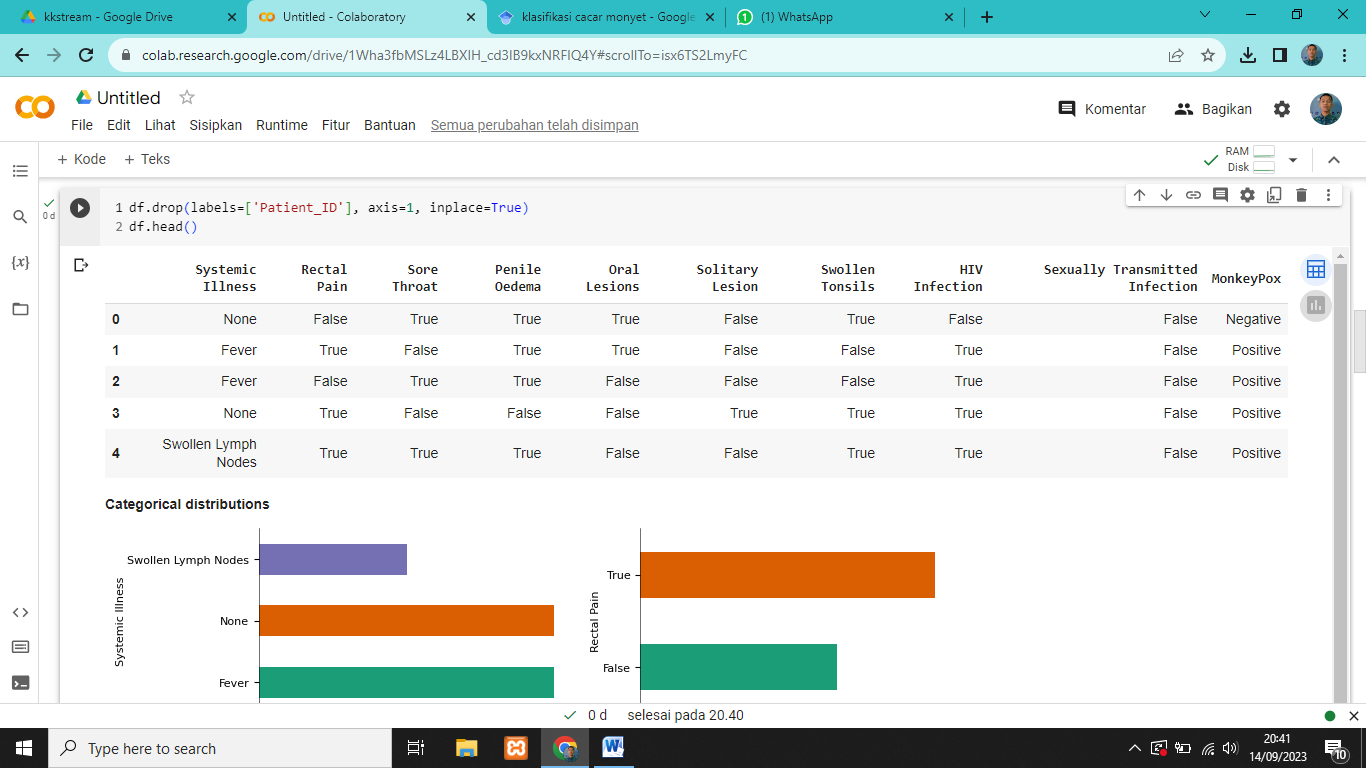
Matkul : Proyek Sains Data E

Tugas Mencari data di kaggle : <https://www.kaggle.com/datasets/muhammad4hmed/monkeypox-patients-dataset>

**Data** tentang penyakit monkeypox(cacar monyet). Data tersebut memiliki total 11 dengan fitur sebanyak 10 kolom sedangkan target 1 kolom dengan jumlah data 25000.

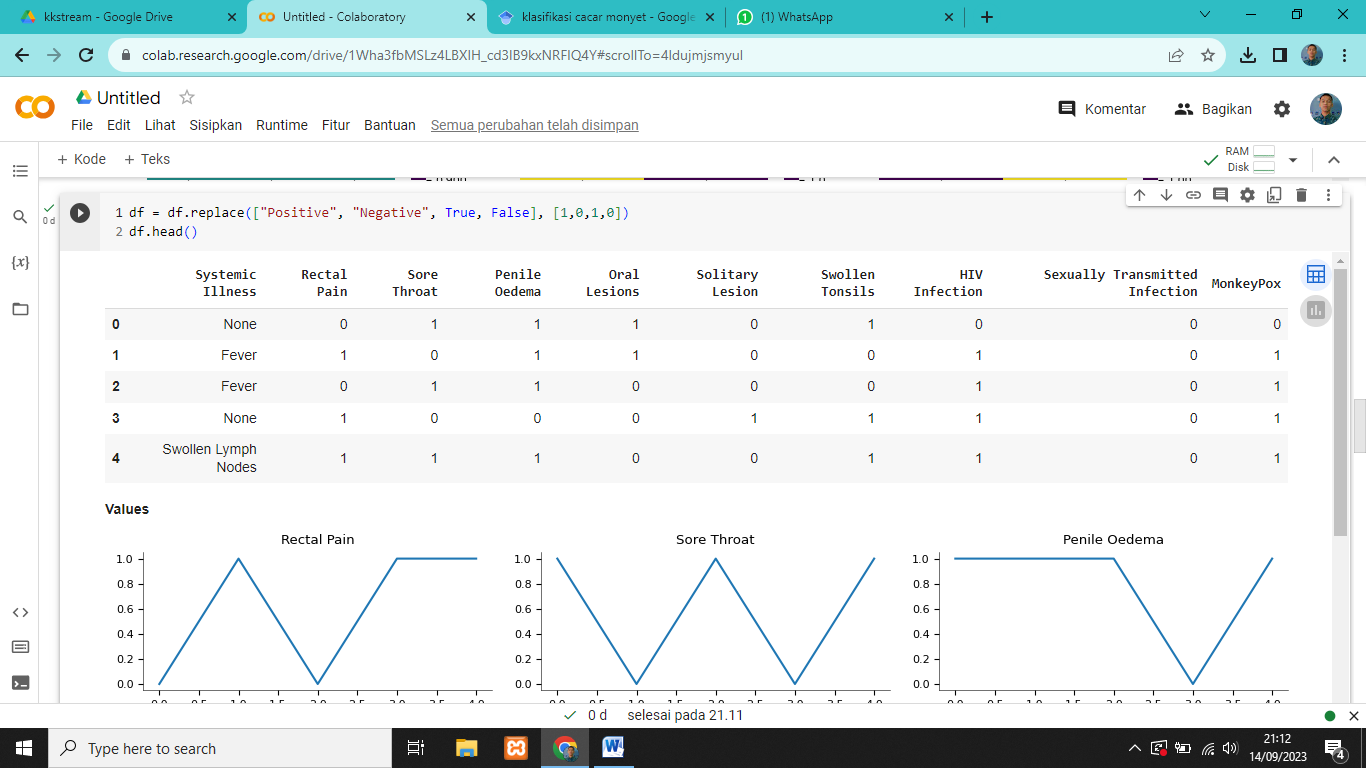


Dari 10 fitur tersebut ada 1 fitur yang nantinya tidak akan digunakan yaitu pada baris Patient\_ID(id pasien). Jadi fitur tersebut akan dihapus karena tidak digunakan. hasilnya seperti ini:

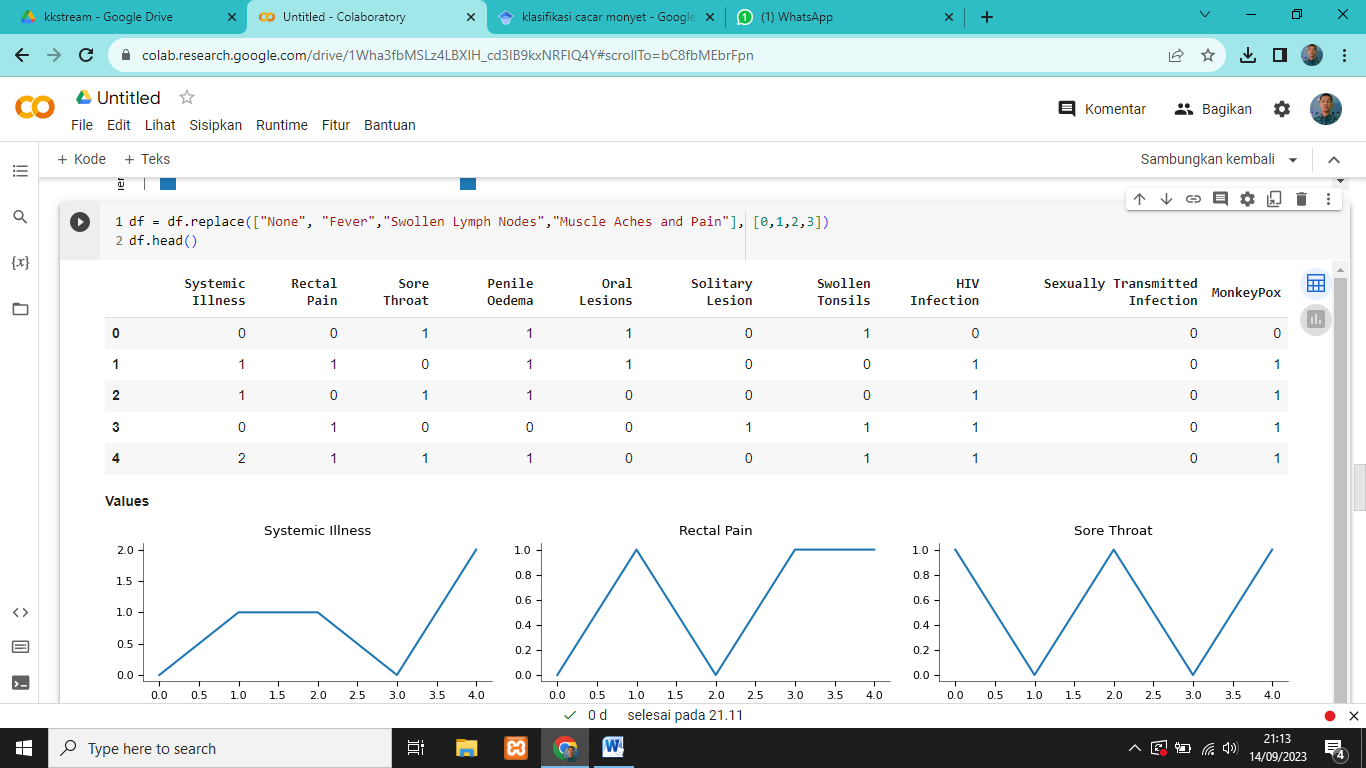


**Teks ke Biner**

Setelah penghapusan maka saya melakukan perubahan dimana data teks akan menjadi biner. Dimana false = 0 dan true = 1, sedangkan untuk kolom pertama itu nilai nya ada 4 yaitu none = 0, fever = 1, swollen lymph nodes 2, dan Muscle Aches and Pain 3. Kenapa data tersebut di ubah karena agar bisa diolah, hasil :

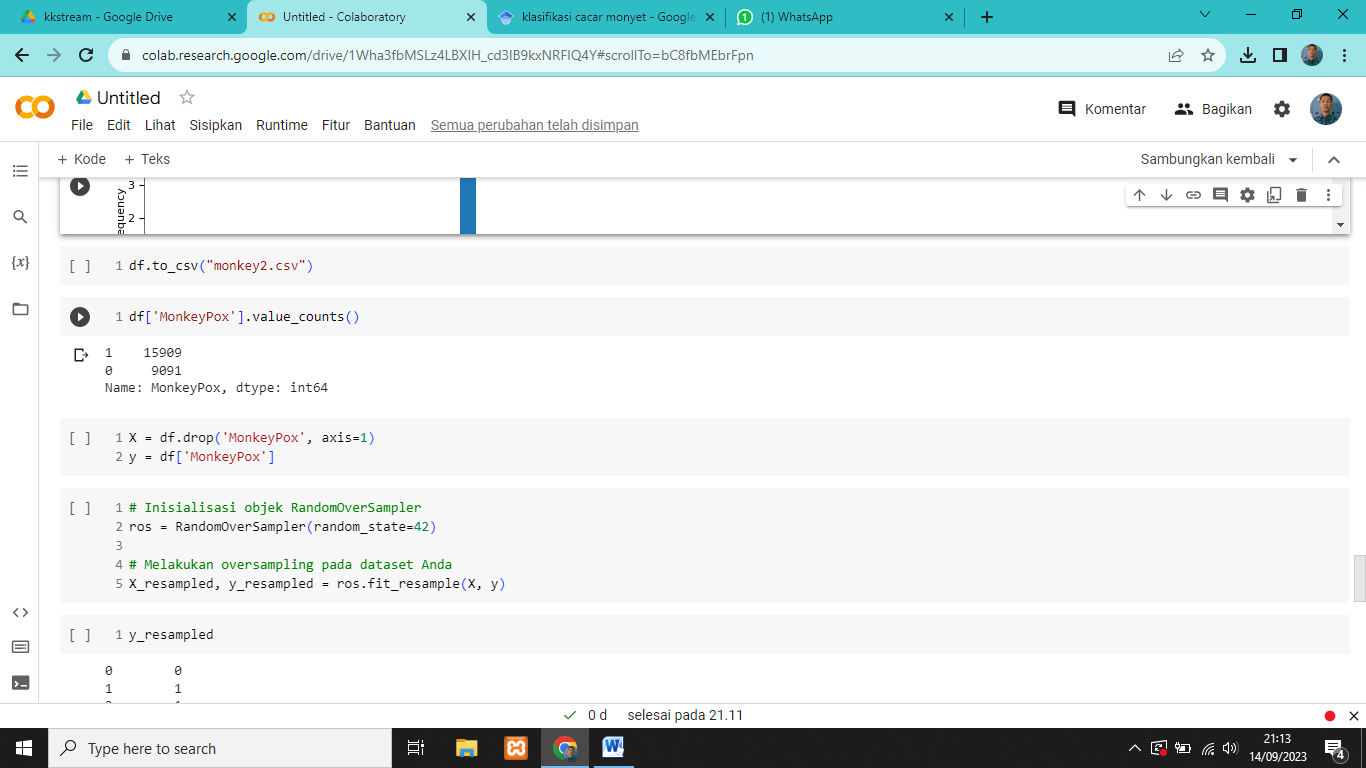


Selanjutnya yaitu di kolom Systematic Illnes merubah ke biner



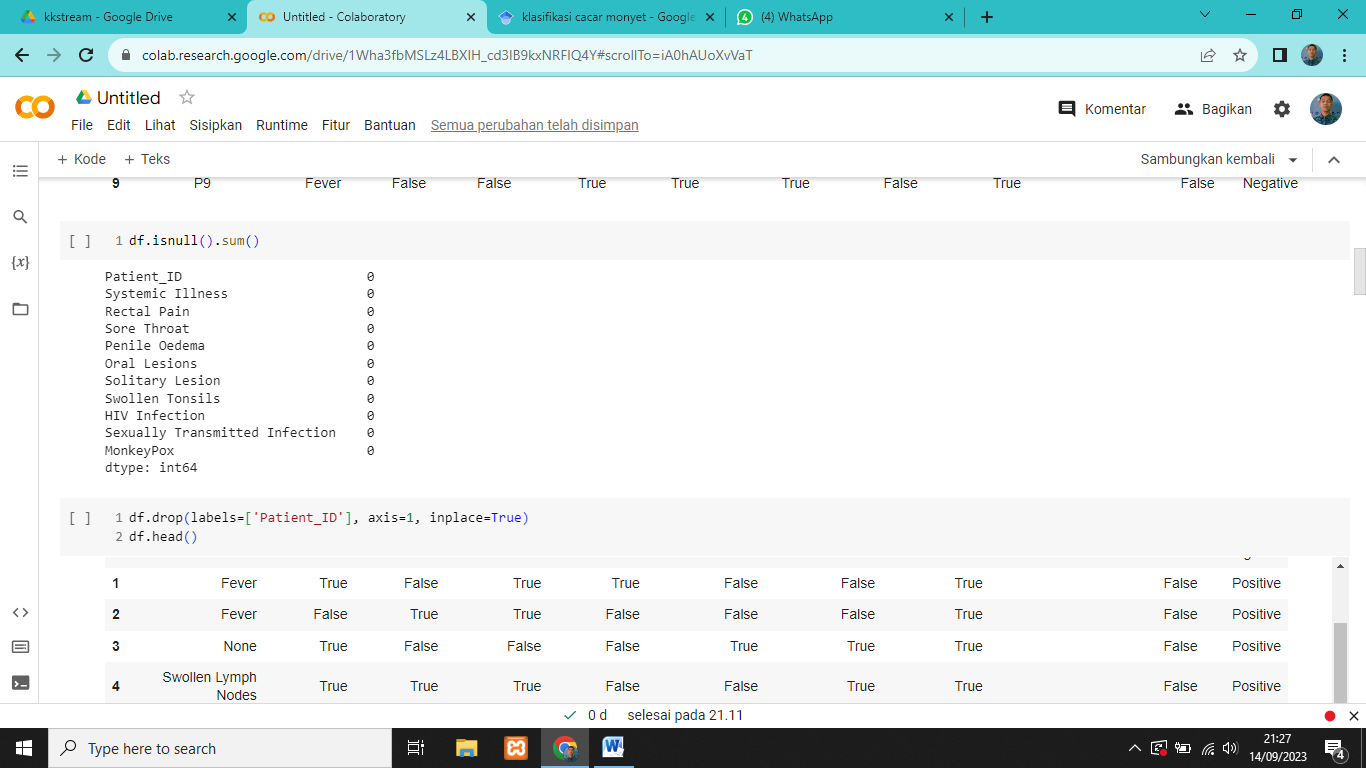
**Data yang tidak seimbang**

Tampilan jumlah data berbeda dan tidak seimbang 15909 untuk positif dan 9091 untuk negative jadi nanti akan saya seimbangkan



Jadi saya menggunakan 9 fitur yaitu pada kolom :

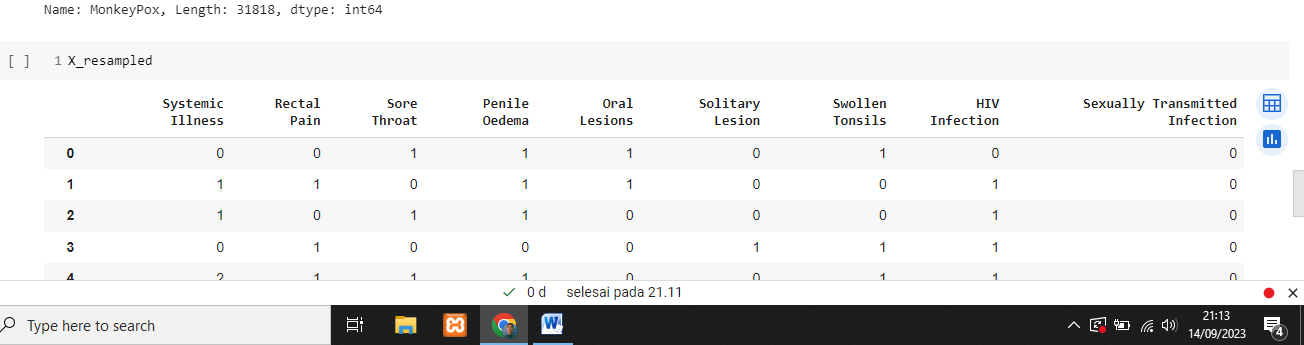
| **Systemic Illness** | **Rectal Pain** | **Sore Throat** | **Penile Oedema** | **Oral Lesions** | **Solitary Lesion** | **Swollen Tonsils** | **HIV Infection** | **Sexually Transmitted Infection** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |



Dan pada masing masing kolom tidak ada data yang missing value atau kosong datanya.

**Balancing Data**

Data yang tadinya hanya 25000 yang positif 15.909 dan negative 9.091 sekarang kita menggunakan ML yang membuat data sedikit akan diperbanyak total data menjadi 31.818 (over sampling)



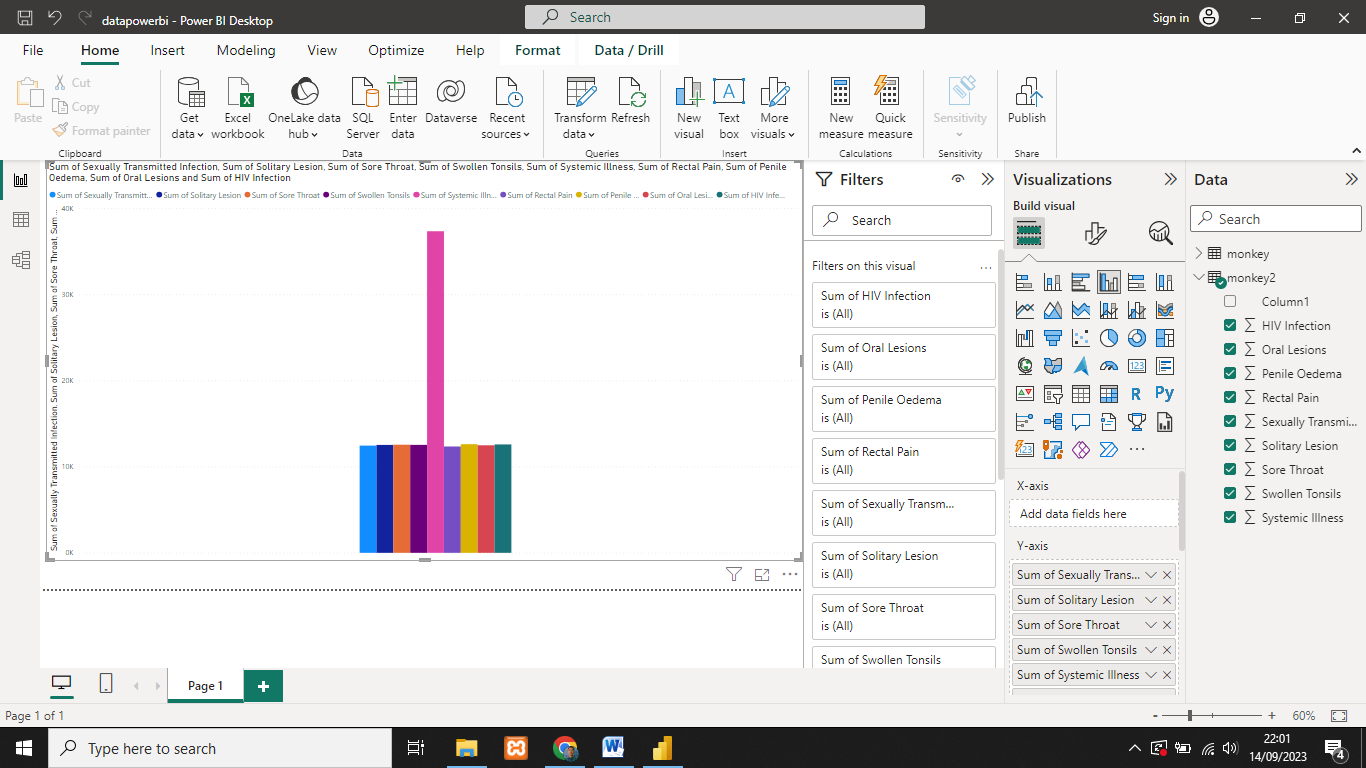
**Pemilihan Fitur**

Penjelasan tentang fitur yang digunakan :

1. **Systemic Illness (Penyakit Sistemik)**: Cacar monyet adalah penyakit sistemik, yang berarti bahwa virus monkeypox dapat menyebar ke seluruh tubuh dan memengaruhi berbagai organ dan sistem. Pasien dengan cacar monyet sering mengalami demam, nyeri tubuh, kelelahan, dan gejala sistemik lainnya.
2. **Rectal Pain (Nyeri Rektal):** Nyeri rektal mungkin bukan gejala khas cacar monyet, dan biasanya lebih terkait dengan masalah pencernaan atau peradangan di daerah rektal daripada dengan infeksi monkeypox itu sendiri.
3. **Sore Throat (Sakit Tenggorokan):** Ini adalah gejala umum yang dapat disebabkan oleh berbagai kondisi, termasuk infeksi virus, infeksi bakteri, atau iritasi.
4. **Penile Edema (Pembengkakan Penis):** Ini merujuk pada pembengkakan atau pembesaran abnormal pada penis. Penyebabnya bisa bervariasi, termasuk infeksi, cedera, atau kondisi medis lainnya.
5. **Oral Lesions (Luka Mulut)**: Monkeypox dapat menyebabkan luka pada kulit, dan dalam beberapa kasus, luka juga dapat terjadi di dalam mulut dan tenggorokan. Ini termasuk lepuh dan lesi yang dapat membuat makan dan minum menjadi sulit.
6. **Solitary Lesion (Luka Tunggal)**: Ini mengacu pada adanya satu luka atau perubahan pada kulit atau dalam tubuh.
7. **Swollen Tonsils (Amandel Bengkak):** Pembengkakan amandel dapat terjadi sebagai respons terhadap infeksi virus, termasuk monkeypox. Amandel yang bengkak dapat menyebabkan nyeri tenggorokan dan kesulitan menelan.
8. **HIV Infection (Infeksi HIV):** Seseorang dengan infeksi HIV yang tidak terkontrol mungkin memiliki sistem kekebalan tubuh yang melemah, sehingga lebih rentan terhadap infeksi, termasuk monkeypox. Oleh karena itu, pasien dengan infeksi HIV dapat mengalami gejala monkeypox yang lebih berat dan berpotensi memiliki prognosis yang lebih buruk.
9. **Sexually Transmitted Infection (Infeksi Menular Seksual)**: Infeksi menular seksual (IMS) seperti HIV dan herpes genital dapat meningkatkan risiko seseorang terinfeksi monkeypox jika mereka memiliki kontak seksual dengan seseorang yang terinfeksi. IMS dapat merusak membran mukosa, yang mungkin mempermudah masuknya virus monkeypox.

**TAMPILAN POWER BI**

Setelah dimasukkan data yang sudah diolah dan dijadikan biner. Lalu dapat dimasukkan ke power BI

****