

**LAPORAN PROJECT
BUSINESS INTELLIGENCE**

Politeknik Caltex Riau

JUDUL

Sentiment Analysis Review Hotel Eropa

Dosen/PLP :

Muhammad Mahrus Zain, S.S.T., M.T.I.

Asmawati, S.Tr.Kom

Identitas :

NO	NAMA	NIM
1	Disya Putri N.	2257301037
2	M. Arif Hamidi	2257301085
3	Riska Oktaviana	2257301120

**PRODI SISTEM INFORMASI
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK CALTEX RIAU**

BAB I

PENDAHULUAN..... 3

- 1.1 Latar Belakang..... 3
- 1.2 Deskripsi Dataset..... 4
- 1.3 Pengambilan Keputusan..... 4

BAB 2

PEMBERSIHAN DATA..... 6

- 2.1 Proses ETL (Extract, Transform, Load)..... 6
- 2.2 Proses 9 Step Kimbal..... 21
 - 1. Memilih Proses Bisnis (Choose the Business Process)..... 21
 - 2. Menentukan Grain (Declare the Grain)..... 21
 - 3. Mengidentifikasi Dimensi (Identify the Dimensions)..... 21
 - 4. Mengidentifikasi Fakta (Identify the Facts)..... 22
 - 5. Menyimpan Pre-Kalkulasi pada Tabel Fakta (Store Precalculated Facts)..... 22
 - 6. Melengkapi Tabel Dimensi (Build the Dimension Tables)..... 22
 - a. Dim_Lokasi..... 22
 - 7. Menentukan Durasi Database (Determine the Duration of the Database)..... 23
 - Rentang Waktu Dataset..... 24
 - 8. Tahapan Preprocessing Data dan ETL (ETL Process and Data Cleaning)..... 24
 - a. Extract..... 24
 - b. Transform..... 24
 - c. Load..... 25
 - 9. Iterasi dan Pengembangan Lanjutan (Deliver and Iterate)..... 25
- 2.3 Perancangan Star Schema..... 25

BAB 3

KESIMPULAN DAN SARAN..... 26

BAB I PENDAHULUAN

1.1 Latar Belakang

Dalam era digital saat ini, ulasan konsumen yang dipublikasikan secara online telah menjadi salah satu faktor utama yang mempengaruhi keputusan calon pelanggan, khususnya dalam industri perhotelan. Platform reservasi hotel, seperti Booking.com dan TripAdvisor, menyediakan ruang bagi tamu hotel untuk memberikan ulasan secara terbuka berdasarkan pengalaman mereka. Ulasan-ulasan ini, yang terdiri dari komentar positif dan negatif, membentuk opini publik terhadap suatu hotel dan dapat berdampak signifikan terhadap reputasi dan tingkat pemesanan hotel tersebut.

Namun, dengan jumlah ulasan yang sangat besar dan terus bertambah setiap harinya, menjadi tidak efisien dan tidak praktis bagi calon pelanggan atau manajemen hotel untuk membaca satu per satu ulasan tersebut secara manual. Oleh karena itu, diperlukan sebuah metode analisis yang dapat secara otomatis mengklasifikasikan dan mengekstraksi sentimen dari data teks dalam jumlah besar. Salah satu pendekatan yang populer dan efektif untuk mengatasi permasalahan ini adalah Sentiment Analysis.

Sentiment Analysis atau analisis sentimen merupakan cabang dari Natural Language Processing (NLP) yang berfokus pada pengolahan dan analisis opini, perasaan, atau sikap pengguna yang diekspresikan dalam bentuk teks. Dengan menerapkan metode ini pada ulasan hotel, informasi berharga seperti kecenderungan opini pelanggan (positif atau negatif), aspek yang paling banyak dikomentari, serta potensi area perbaikan dapat dievaluasi secara objektif dan efisien.

Proyek ini menggunakan dataset berjudul "515k Hotel Reviews Data in Europe" yang diperoleh dari platform Kaggle. Dataset ini berisi lebih dari 500 ribu ulasan hotel yang dikumpulkan dari berbagai negara di Eropa, termasuk teks ulasan serta label sentimen berupa skor positif dan negatif. Melalui proses pra-pemrosesan data, ekstraksi fitur, dan penerapan algoritma machine learning atau deep learning, penelitian ini bertujuan untuk membangun model yang mampu mengklasifikasikan sentimen ulasan secara otomatis dan akurat.

1.2 Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini berjudul "515k Hotel Reviews Data in Europe" dan diambil dari platform Kaggle. Dataset ini terdiri dari lebih dari 500.000 ulasan hotel yang dikumpulkan dari berbagai negara di Eropa, seperti Inggris, Prancis, Italia, Jerman, dan negara-negara lainnya. Data ini diperoleh dari pengguna situs pemesanan hotel, seperti Booking.com, dan mencakup berbagai elemen penting yang dapat digunakan untuk analisis sentimen.

Beberapa fitur utama dalam dataset ini antara lain:

- Hotel Name: Nama hotel tempat ulasan diberikan.
- Reviewer Nationality: Kebangsaan dari pengguna yang memberikan ulasan.
- Review Date: Tanggal ulasan dibuat.
- Positive Review: Kolom yang berisi komentar positif dari pelanggan.
- Negative Review: Kolom yang berisi komentar negatif dari pelanggan.
- Reviewer Score: Skor yang diberikan oleh pelanggan (dalam skala 1–10).
- Total Number of Reviews: Jumlah total ulasan yang diberikan oleh pengguna tersebut.
- Average Score: Nilai rata-rata hotel berdasarkan seluruh ulasan.
- Tags: Informasi tambahan terkait perjalanan pengguna, seperti apakah perjalanan dilakukan untuk bisnis atau liburan, jumlah malam menginap, dan jenis kamar.

1.3 Pengambilan Keputusan

Analisis sentimen yang dilakukan dalam penelitian ini bertujuan untuk mendukung proses pengambilan keputusan secara lebih objektif dan efisien, baik dari sisi konsumen maupun manajemen hotel. Dengan mengklasifikasikan ulasan pelanggan ke dalam kategori sentimen positif dan negatif, informasi yang dihasilkan dapat digunakan untuk berbagai keperluan strategis, antara lain:

1. Bagi Konsumen

Memberikan ringkasan opini secara otomatis terhadap hotel-hotel yang ada, sehingga memudahkan calon pelanggan dalam memilih hotel yang sesuai dengan preferensi dan ekspektasi mereka tanpa harus membaca semua ulasan satu per satu.

2. Bagi Manajemen Hotel

Memberikan wawasan terkait kelebihan dan kekurangan layanan yang dirasakan oleh pelanggan. Ulasan negatif dapat digunakan sebagai dasar untuk melakukan evaluasi dan peningkatan layanan, sementara ulasan positif dapat menjadi indikator kekuatan kompetitif yang perlu dipertahankan.

3. Bagi Peneliti atau Pengembang Sistem

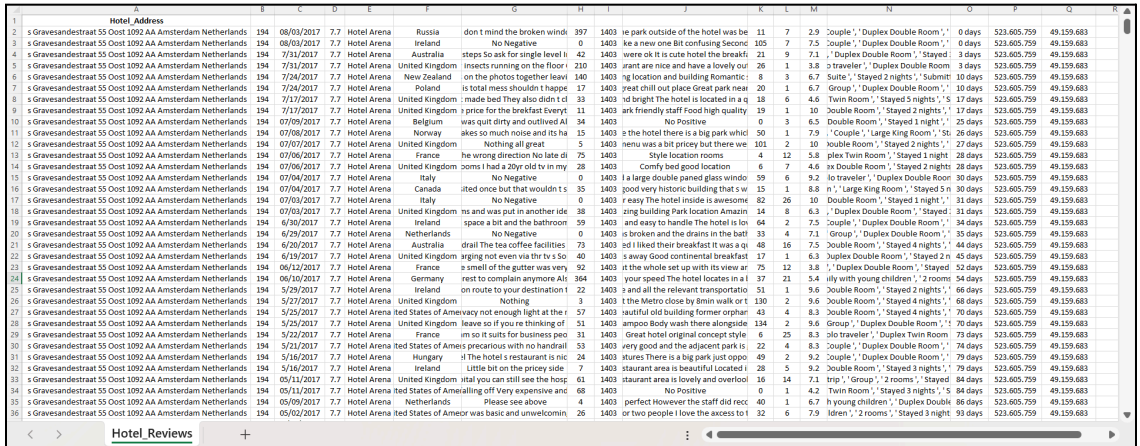
Hasil dari model analisis sentimen ini dapat diterapkan dalam sistem rekomendasi, chatbot customer service, ataupun sistem peringatan dini terhadap penurunan kualitas layanan di hotel tertentu. Dengan kata lain, analisis sentimen terhadap ulasan hotel mampu membantu dalam proses pengambilan keputusan yang lebih cepat, tepat, dan berbasis data (data-driven decision making), yang sangat dibutuhkan di era digital seperti saat ini.

BAB 2

PEMBERSIHAN DATA

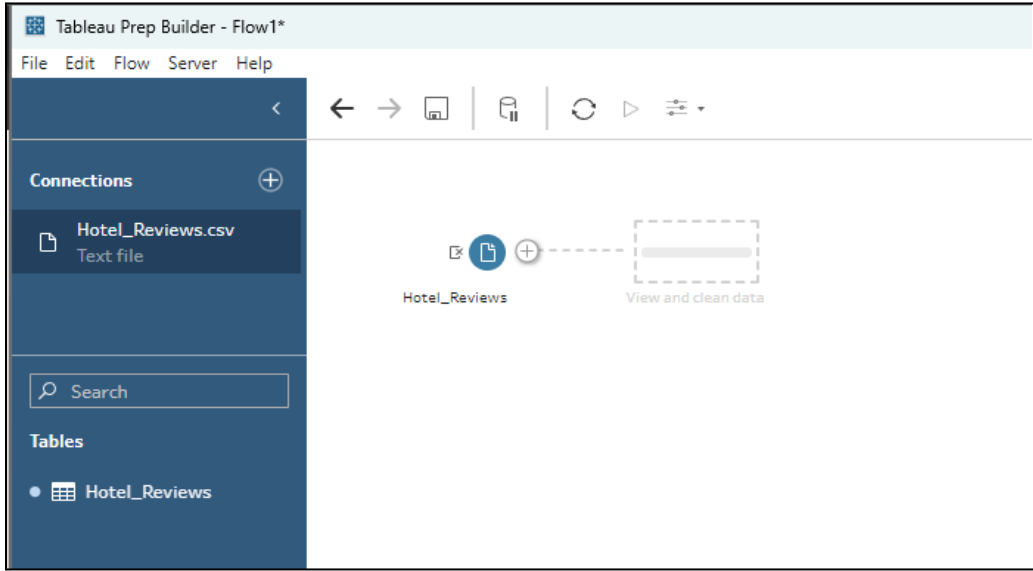
2.1 Proses ETL (Extract, Transform, Load)

Proses analisis sentimen terhadap ulasan hotel memerlukan tahapan awal untuk mempersiapkan data agar dapat diolah secara optimal. Salah satu tahapan penting dalam pengelolaan data adalah ETL (Extract, Transform, Load), yaitu rangkaian proses yang mencakup pengumpulan data (extract), pembersihan dan transformasi data (transform), serta pemuatan data ke dalam format yang siap digunakan (load). Melalui proses ETL, kualitas dan konsistensi data dapat dijaga sehingga mendukung hasil analisis yang lebih akurat dan bermakna. Berikut proses ETL ditunjukkan pada tabel di bawah ini yang merangkum tahapan pengambilan, transformasi, dan pemuatan data dalam penelitian analisis sentimen ulasan hotel Eropa:

NO	Dokumentasi	
1.	<p>DATASET</p>  <p>The screenshot displays a dataset of hotel reviews. The columns are: Hotel Address, Date, Rating, Country, Review, and Sentiment. The data includes reviews from various hotels in different countries, with ratings ranging from 1 to 5 and sentiment labels like 'Positive', 'Negative', and 'Neutral'.</p>	
	<p>Analisis</p> <p>Sebelum masuk ke tahap analisis, peneliti terlebih dahulu melakukan proses preprocessing data di Tableau Prep. Tahapan ini bertujuan untuk memastikan data bersih dari nilai null atau kosong, sehingga proses analisis dapat berjalan dengan lebih</p>	

Business Intelligence

akurat dan tanpa gangguan.

NO	Dokumentasi
2.	 <p>The screenshot shows the Tableau Prep Builder interface. On the left, there is a sidebar with 'Connections' and 'Tables'. Under 'Connections', 'Hotel_Reviews.csv' (Text file) is listed. Under 'Tables', 'Hotel_Reviews' is listed. The main workspace shows a flow diagram with a connection icon and a table icon, labeled 'Hotel_Reviews', and a dashed box labeled 'View and clean data'.</p>
	Analisis
	<p>Langkah awal dimulai dengan menghubungkan Tableau Prep ke dataset yang akan diproses dalam tahapan ETL. Proses ini dilakukan dengan memilih file data yang akan digunakan, sehingga Tableau Prep dapat membaca dan menampilkan isi dataset untuk langkah transformasi selanjutnya.</p>

Business Intelligence

NO

Dokumentasi

3.

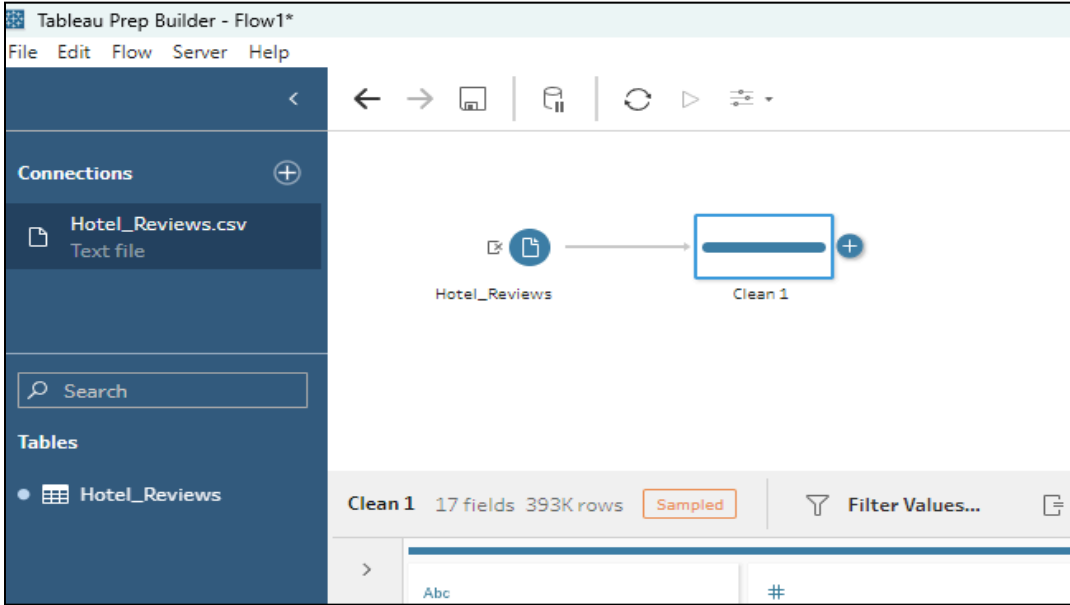
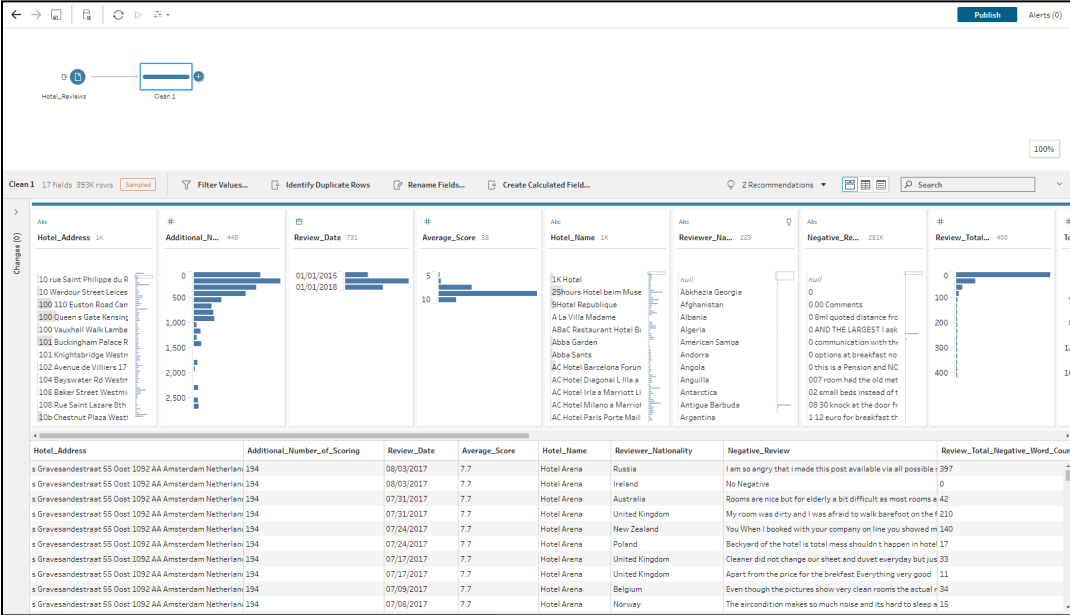
The screenshot shows the Tableau Prep Builder interface. On the left, the 'Connections' pane lists 'Hotel_Reviews.csv' as a 'Text File'. Below it, the 'Tables' pane shows 'Hotel_Reviews'. The main area is divided into 'Settings', 'Tables', 'Data Sample', and 'Changes (1)'. The 'Settings' pane on the left shows 'Connection Details' (Text File, File path: C:\Users\j...Downloads\Hotel_Reviews.csv), 'Header Options' (Generate header (in F1, F2, F3) is selected), and 'Text Options' (Field Separator: Comma, Text Qualifier: Automatic, Character Set: UTF-8). The 'Preview' tab on the right shows a table with 18 fields: Row ID, Hotel_Address, Additional_Number_of_Suited, Review_Date, Average_Score, Hotel_Name, and Reviewer_Name. The table contains 21 rows of data.

Row ID	Hotel_Address	Additional_Number_of_Suited	Review_Date	Average_Score	Hotel_Name	Reviewer_Name
1	s Gravesandestrat S...	194	06/03/2017	7.7	Hotel Arena	Russia
2	s Gravesandestrat S...	194	06/03/2017	7.7	Hotel Arena	Iranian
3	s Gravesandestrat S...	194	07/30/2017	7.7	Hotel Arena	Austria
4	s Gravesandestrat S...	194	07/30/2017	7.7	Hotel Arena	United
5	s Gravesandestrat S...	194	07/04/2017	7.7	Hotel Arena	Hevi 2
6	s Gravesandestrat S...	194	07/04/2017	7.7	Hotel Arena	Poland
7	s Gravesandestrat S...	194	07/17/2017	7.7	Hotel Arena	United
8	s Gravesandestrat S...	194	07/17/2017	7.7	Hotel Arena	United
9	s Gravesandestrat S...	194	07/09/2017	7.7	Hotel Arena	Belgium
10	s Gravesandestrat S...	194	07/06/2017	7.7	Hotel Arena	Norway
11	s Gravesandestrat S...	194	07/07/2017	7.7	Hotel Arena	United
12	s Gravesandestrat S...	194	07/06/2017	7.7	Hotel Arena	France
13	s Gravesandestrat S...	194	07/06/2017	7.7	Hotel Arena	United
14	s Gravesandestrat S...	194	07/04/2017	7.7	Hotel Arena	Italy
15	s Gravesandestrat S...	194	07/04/2017	7.7	Hotel Arena	Canada
16	s Gravesandestrat S...	194	07/03/2017	7.7	Hotel Arena	Italy
17	s Gravesandestrat S...	194	07/03/2017	7.7	Hotel Arena	United
18	s Gravesandestrat S...	194	06/30/2017	7.7	Hotel Arena	Iranian
19	s Gravesandestrat S...	194	06/29/2017	7.7	Hotel Arena	Netherlands
20	s Gravesandestrat S...	194	06/20/2017	7.7	Hotel Arena	Austria
21	s Gravesandestrat S...	194	06/20/2017	7.7	Hotel Arena	Austria

Analisis

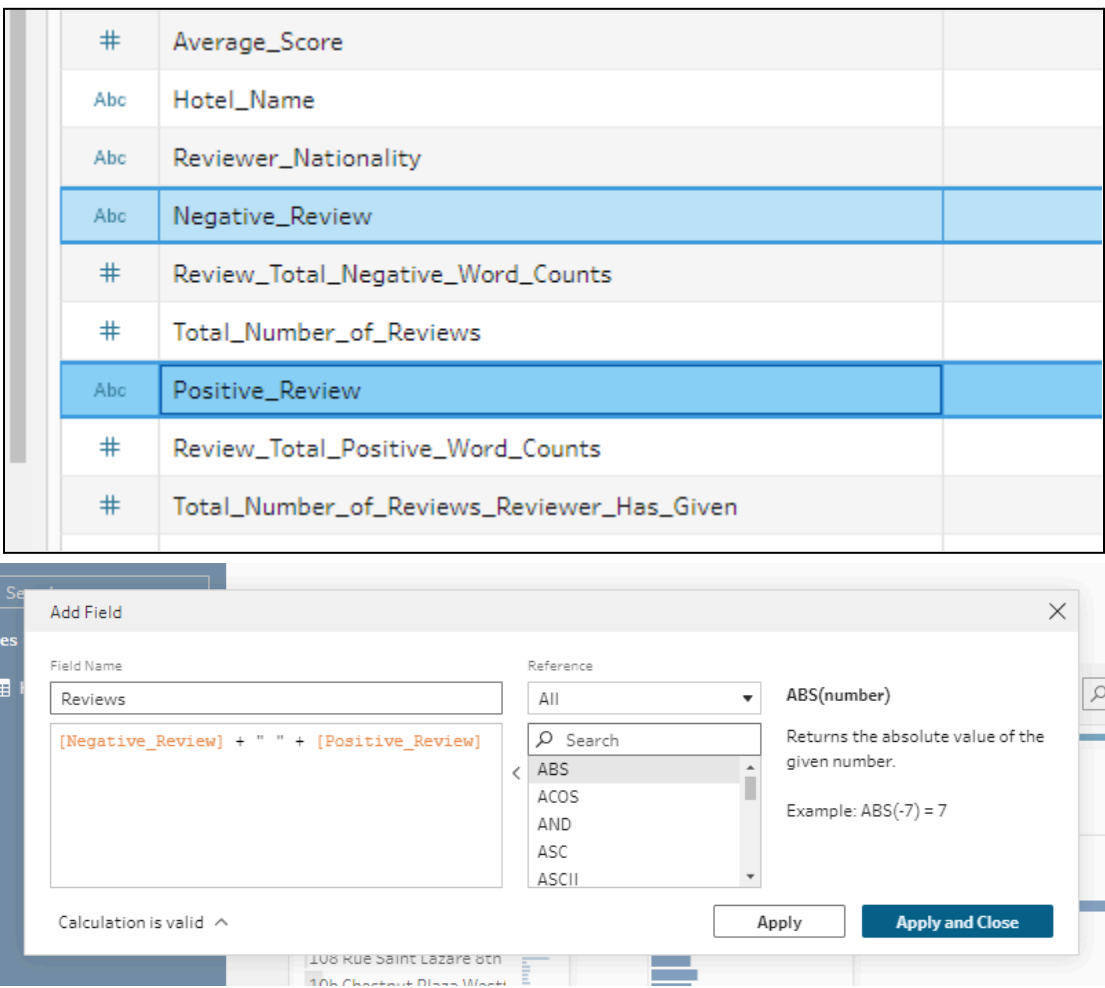
Setelah halaman input terbuka, peneliti melakukan pengaturan awal terhadap data yang dimuat. Langkah ini mencakup menjadikan baris pertama sebagai nama kolom (header) dan memastikan bahwa data dipisahkan menggunakan tanda koma (,) sebagai pemisah antar nilai. Pengaturan ini penting untuk memastikan struktur data terbaca dengan benar di Tableau Prep.

Business Intelligence

NO	Dokumentasi
4.	 <p>The screenshot shows the Tableau Prep Builder interface. On the left, the 'Connections' pane lists 'Hotel_Reviews.csv' as a 'Text file'. Below it, the 'Tables' pane shows 'Hotel_Reviews'. The main workspace displays a flow: 'Hotel_Reviews' (represented by a document icon) connects to a 'Clean 1' step (represented by a blue box with a plus sign). Below the flow, a summary bar indicates 'Clean 1' has '17 fields', '393K rows', and is 'Sampled'. At the bottom, a preview of the data is visible, showing columns like 'Hotel_Address', 'Additional_Number_of_Scoring', 'Review_Date', 'Average_Score', 'Hotel_Name', 'Reviewer_Nationality', 'Negative_Review', and 'Review_Total_Negative_Word_Counts'.</p>  <p>The second screenshot shows the 'Clean 1' step in detail. It displays a data preview with columns: 'Hotel_Address', 'Additional_Number_of_Scoring', 'Review_Date', 'Average_Score', 'Hotel_Name', 'Reviewer_Nationality', 'Negative_Review', and 'Review_Total_Negative_Word_Counts'. The preview shows data for 'Hotel_Arena' in 'Russia' and 'Ireland', with negative reviews mentioning 'No Negative', 'Rooms are nice but for elderly a bit difficult as most rooms are 42', 'My room was dirty and I was afraid to walk barefoot on the floor', 'You When I booked with your company on line you showed me 140', 'Backyard of the hotel is total mess shouldn't happen in hotel', 'Cleaner did not change our sheet and duvet everyday but just 33', 'Apart from the price for the breakfast Everything very good', 'Even though the pictures show very clean rooms the actual', and 'The aircondition makes so much noise and its hard to sleep'.</p>

Business Intelligence

Analysis	
	Langkah berikutnya, peneliti mulai melakukan proses transformasi data dengan mengklik ikon tanda plus (+) pada alur (<i>flow</i>), kemudian memilih opsi Clean Step. Fitur ini digunakan untuk membersihkan data, seperti menggabungkan kolom, memisahkan nilai dalam kolom, serta melakukan penyesuaian lain yang dibutuhkan sebelum data dianalisis lebih lanjut.

NO	Dokumentasi
5.	 <p>The screenshot displays a data transformation interface. At the top, a list of fields is shown with a search bar on the left. The fields are: Average_Score, Hotel_Name, Reviewer_Nationality, Negative_Review, Review_Total_Negative_Word_Counts, Total_Number_of_Reviews, Positive_Review, Review_Total_Positive_Word_Counts, and Total_Number_of_Reviews_Reviewer_Has_Given. Below this list, an 'Add Field' dialog box is open. The dialog box has a 'Field Name' field containing 'Reviews' and a 'Reference' dropdown set to 'All'. The formula '[Negative_Review] + " " + [Positive_Review]' is entered in the formula field. A list of functions is shown, with 'ABS' selected. The description for 'ABS(number)' is 'Returns the absolute value of the given number. Example: ABS(-7) = 7'. The dialog box also includes 'Apply' and 'Apply and Close' buttons.</p>

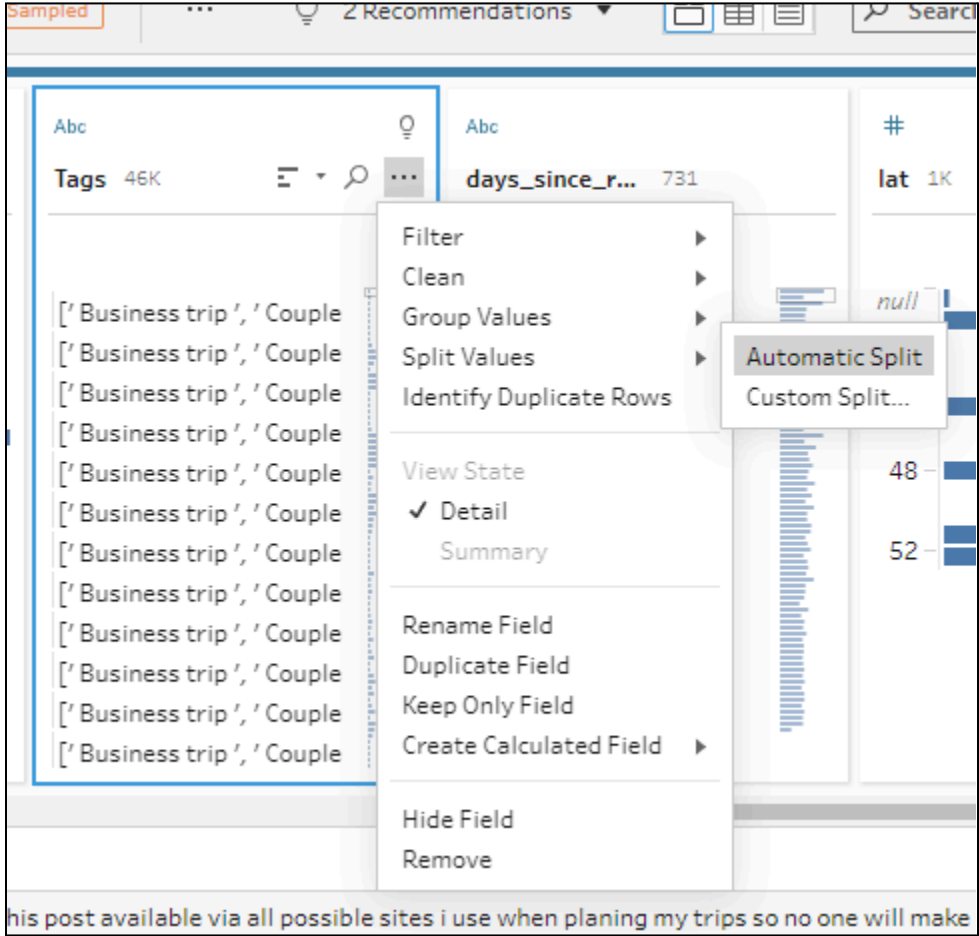
Business Intelligence

	<div data-bbox="300 325 1404 619"> <p>Reviews</p> <p>You When I booked with your company on line you showed me pictures of a room I thought I was getting and paying for and then when we arrived that s room was booked and the staff told me we could s Gri</p> <p>Backyard of the hotel is total mess should n t happen in hotel with 4 stars Good restaurant with modern design great chill out place Great park nearby the hotel and awesome main stairs s Gri</p> <p>Cleaner did not change our sheet and duvet everyday but just made bed They also didn t clean the floor and changed the body gel when we run out of it The room is spacious and bright The hotel is locat s Gri</p> <p>Apart from the price for the breakfast Everything very good Good location Set in a lovely park friendly staff Food high quality We Oth enjoyed the breakfast s Gri</p> <p>Even though the pictures show very clean rooms the actual room was quit dirty and outlived Also check in is at 15 o clock but our room was not ready at that time No Positive s Gri</p> <p>The aircondition makes so much noise and its hard to sleep at night The room was big enough and the bed is good The breakfast food and service on the hotel is good outside the hotel there is a big park s Gri</p> <p>Nothing all great Rooms were stunningly decorated and really spacious in the top of the building Pictures are of room 300 The true beauty of the building has been kept but modernised brilliantly Also s Gri</p> <p>6 30 AM started big noise workers loading wood down the windows Stupid room numbering system It took 20 Minutes with a night guard finally to find our rooms The check in staff pointed us in the wr s Gri</p> <p>The floor in my room was filthy dirty Very basic rooms I had a 20yr old tv in my room Fridge did not work Overpriced breakfast Comfy bed good location s Gri</p> <p>No Negative This hotel is being renovated with great care and with an appreciation for its unique structure and location My spacious and comfortable room had a large double paned glass window onto s Gri</p> </div> <div data-bbox="795 640 909 682"> <p>Analisis</p> </div> <div data-bbox="292 703 1412 913"> <p>Langkah selanjutnya adalah menggabungkan kolom Negative Review dan Positive Review menjadi satu kolom, yang disebut proses merge. Tujuannya agar semua ulasan terkumpul dalam satu kolom, sehingga lebih mudah diolah untuk analisis sentimen. Setelah digabung, kolom tersebut diberi nama baru yaitu Reviews.</p> </div>
--	---

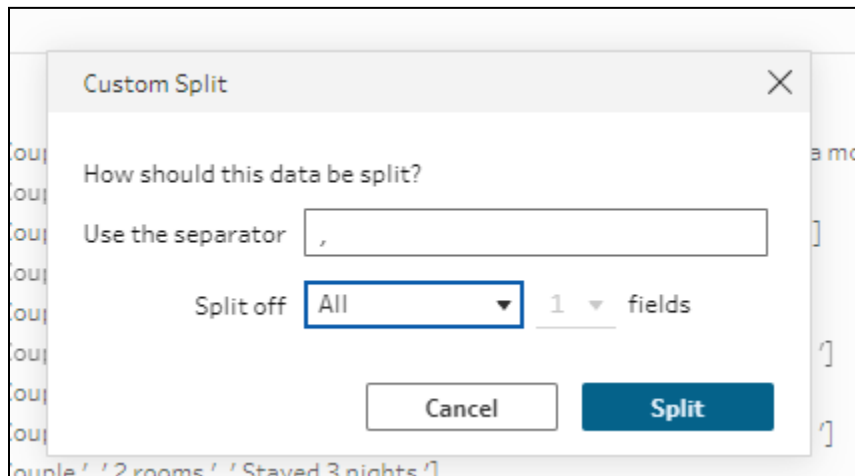
NO	Dokumentasi
6.	<p>Selanjutnya, pada data Reviews ini terdapat null data. Null data ini akan dihapus dengan cara, sebagai berikut:</p> <div data-bbox="641 1186 1079 1837"> </div> <p>Exclude:</p>

Business Intelligence

	Analisis
	Selanjutnya, peneliti menambahkan aktivitas untuk melakukan <i>split</i> pada kolom <i>Tags</i> . Langkah ini dilakukan karena data pada kolom tersebut masih berisi informasi yang tidak terstruktur dan sulit dibaca. Oleh karena itu, kolom <i>Tags</i> perlu dipecah menjadi beberapa bagian agar lebih rapi dan mudah dianalisis.

NO	Dokumentasi
7.	<p>Kemudian menambah sebuah activity yaitu untuk melakukan Split Tags, Hal ini dilakukan karena data pada field Tags masih sangat berantakan dan tidak beraturan. Sehingga peneliti memilih untuk melakukan split field tags</p>  <p>The screenshot shows a data table with columns 'Tags', 'days_since_r...', and 'lat'. The 'Tags' column contains a list of values, each starting with '[' Business trip ', ' Couple'. A context menu is open over the 'Tags' column, and the 'Split Values' option is selected, showing a sub-menu with 'Automatic Split' and 'Custom Split...'.</p>

Business Intelligence

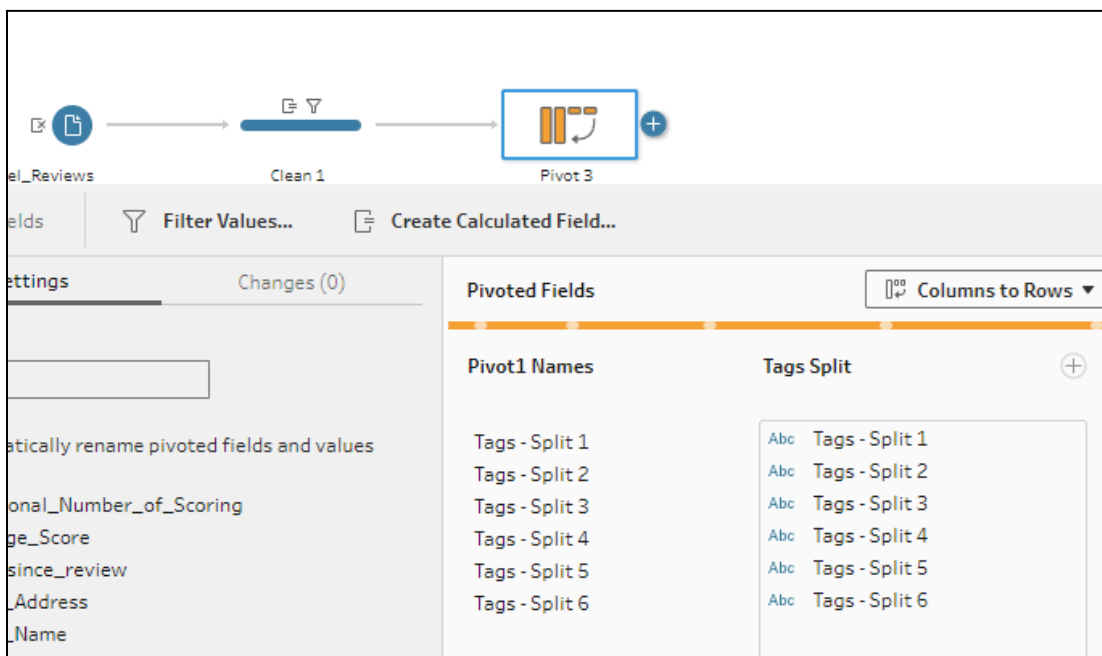


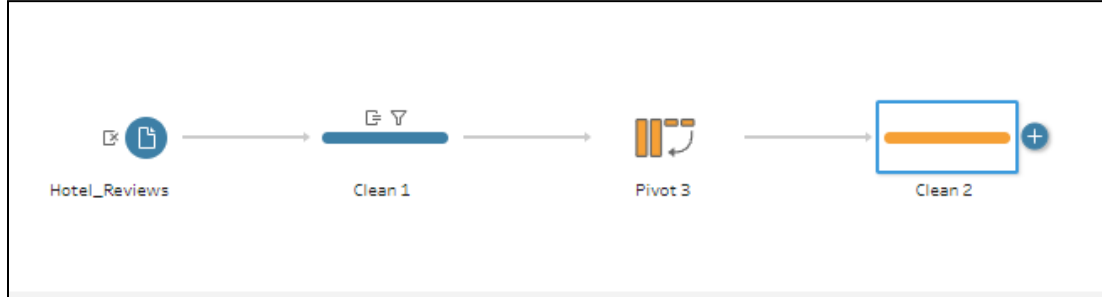
Setelah di split, muncul kolom sebagai berikut:


Analisis

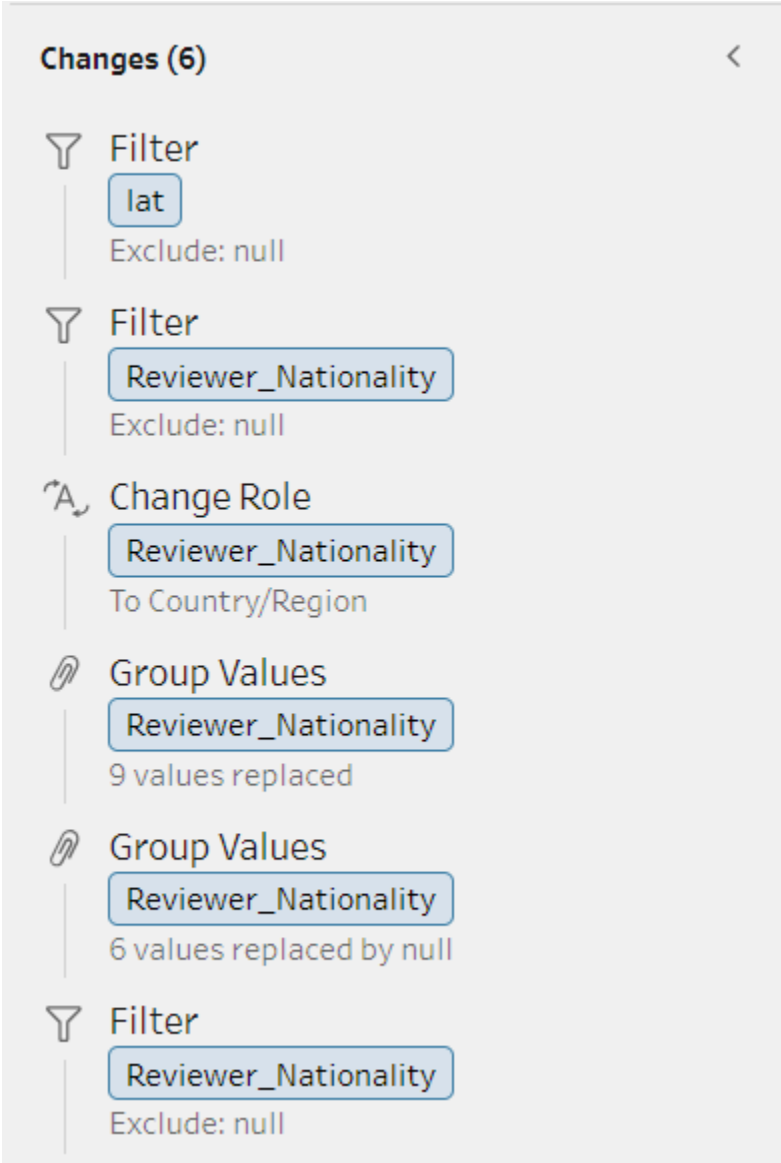
NO	Dokumentasi
8.	Dikarenakan column diatas tidak terlihat baik untuk diolah, maka akan kita lakukan pivot table. Pivot ini berguna untuk mengubah kolom menjadi baris atau sebaliknya

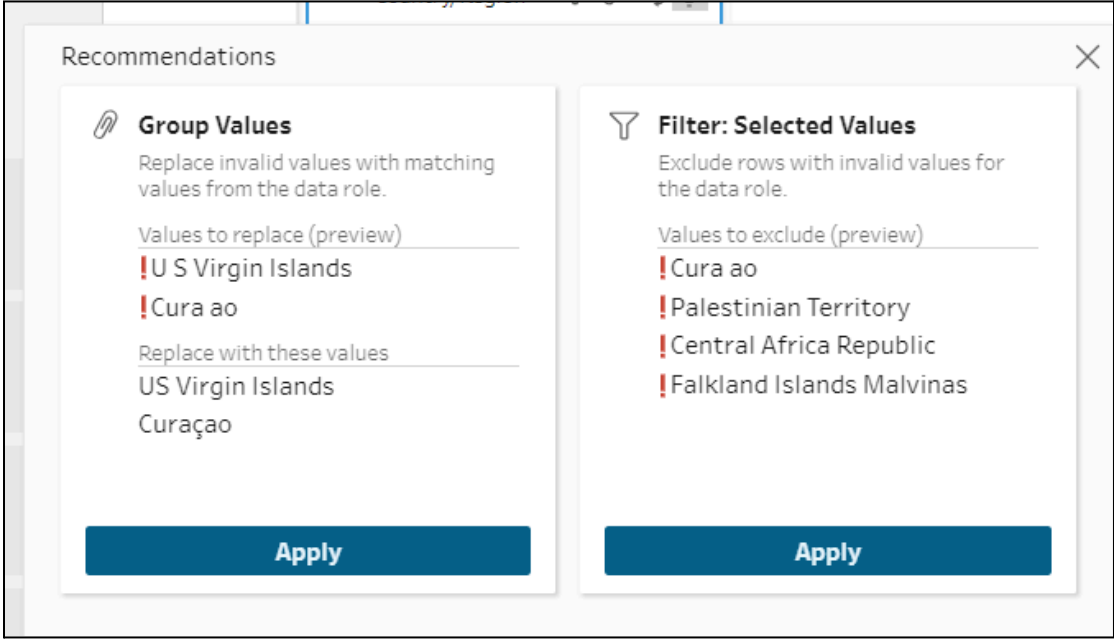
Business Intelligence

	 <p>Hotel_Reviews → Clean 1 → Pivot 3</p> <p>Fields Filter Values... Create Calculated Field...</p> <p>Settings Changes (0)</p> <p>Pivoted Fields Columns to Rows</p> <p>Pivot1 Names Tags Split</p> <p>Tags - Split 1 Tags - Split 1</p> <p>Tags - Split 2 Tags - Split 2</p> <p>Tags - Split 3 Tags - Split 3</p> <p>Tags - Split 4 Tags - Split 4</p> <p>Tags - Split 5 Tags - Split 5</p> <p>Tags - Split 6 Tags - Split 6</p>
	<p style="text-align: center;">Analisis</p>

NO	<p style="text-align: center;">Dokumentasi</p>
9.	<p>Kemudian menambah activity baru untuk melakukan penghapusan data-data yang masih null dari masing-masing field.</p>  <p>Hotel_Reviews → Clean 1 → Pivot 3 → Clean 2</p>
	<p style="text-align: center;">Analisis</p>

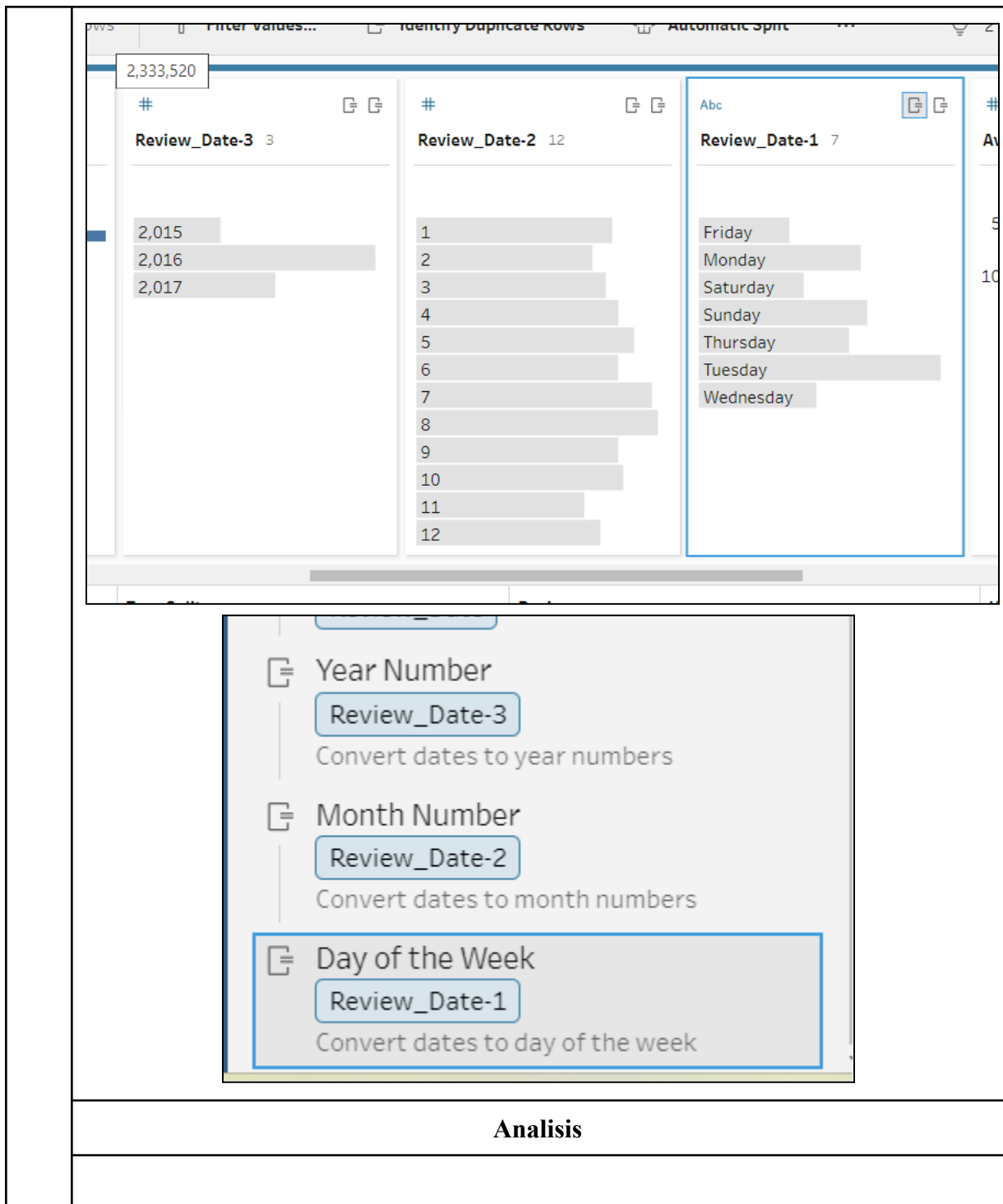
NO	Dokumentasi
10.	<p>Berikut perubahan-perubahan yang dilakukan:</p> 
	<p style="text-align: center;">Analisis</p> <p>Perubahan-perubahan di samping dilakukan untuk menghilangkan nilai-nilai null dan menghapus beberapa field yang tidak bagus untuk diolah, seperti 0 dan 0 00 comments pada field Reviews.</p>

NO	Dokumentasi
11.	<p>Kemudian peneliti melakukan pengubahan tipe data dari fields Reviewer_Nationality. Hal ini dilakukan agar nantinya mudah untuk dilakukan EDA</p>  <p>The screenshot displays a list of six data transformation steps for the 'Reviewer_Nationality' field:</p> <ul style="list-style-type: none"> Filter: Filter by 'lat', excluding null values. Filter: Filter by 'Reviewer_Nationality', excluding null values. Change Role: Change the role of 'Reviewer_Nationality' to 'Country/Region'. Group Values: Group values for 'Reviewer_Nationality', with 9 values replaced. Group Values: Group values for 'Reviewer_Nationality', with 6 values replaced by null. Filter: Filter by 'Reviewer_Nationality', excluding null values.
	<p style="text-align: center;">Analisis</p>

NO	Dokumentasi
12.	<p>Selain mengubah tipe data, peneliti juga mengganti beberapa daerah yang terdapat typo di dalamnya seperti gambar dibawah ini.</p>  <p>The screenshot shows a 'Recommendations' dialog box with two main sections. The left section, titled 'Group Values', contains a description 'Replace invalid values with matching values from the data role.', a list of 'Values to replace (preview)' including 'U S Virgin Islands' and 'Cura ao', and a list of 'Replace with these values' including 'US Virgin Islands' and 'Curaçao'. The right section, titled 'Filter: Selected Values', contains a description 'Exclude rows with invalid values for the data role.', a list of 'Values to exclude (preview)' including 'Cura ao', 'Palestinian Territory', 'Central Africa Republic', and 'Falkland Islands Malvinas'. Both sections have an 'Apply' button at the bottom.</p>
	Analisis

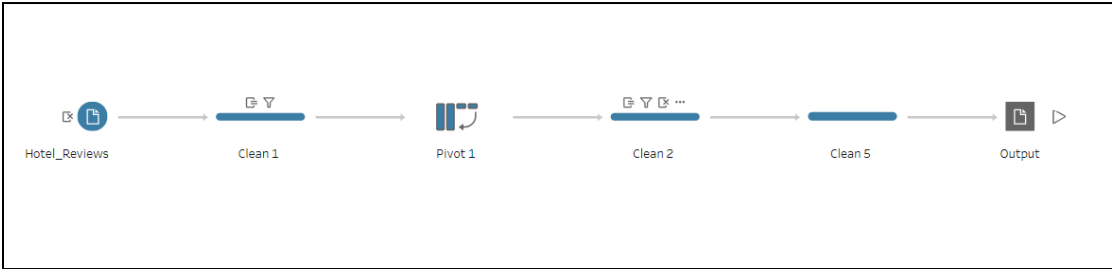
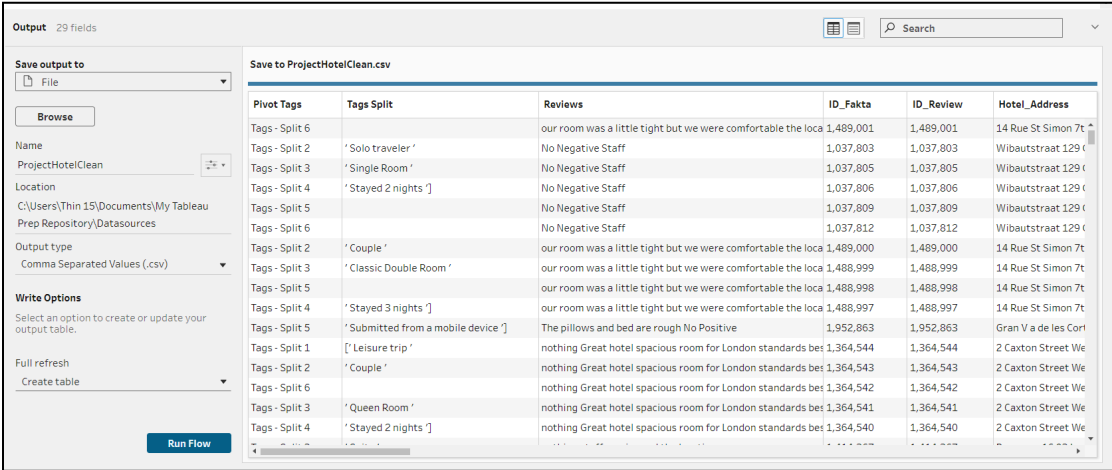
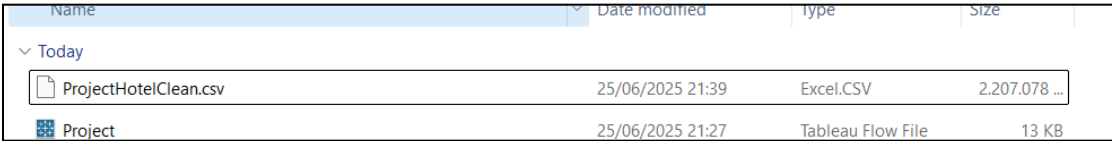
NO	Dokumentasi
13.	Selanjutnya menambahkan beberapa kolom baru untuk dimensi waktu, yaitu days, months, serta years.

Business Intelligence



NO	Dokumentasi
14.	<p>Kemudian membuat masing-masing id untuk masing-masing dimensi yang akan dibutuhkan. Kita dapat menggunakan query di create calculate field dan masukkan query berikut <code>//{ ORDERBY [yourfield] ASC: ROW_NUMBER() }</code> Sehingga akan muncul perubahan sebagai berikut:</p> <p>The screenshot shows a list of fields in a data tool. There are four 'Calculated Field' entries and one 'Rename Field' entry. Each calculated field has a name in a blue box and a query below it. The rename field has a name in a blue box and a description of the change.</p> <ul style="list-style-type: none"> Calculated Field ID_Kota { ORDERBY [Hotel_Address] ASC: ROW_NUMBER() } Calculated Field ID_Reviewer { ORDERBY [Reviewer_Nationality] ASC: ROW_NUMBER() } Calculated Field ID_Hotel { ORDERBY [Hotel_Name] ASC: ROW_NUMBER() } Calculated Field ID_Review { ORDERBY [Reviews] ASC: ROW_NUMBER() } Rename Field ID_Alamat From [ID_Kota] to [ID_Alamat]
	<p style="text-align: center;">Analisis</p>

Business Intelligence

NO	Dokumentasi																																																																																																																						
15.	<p>Kemudian melakukan cleaning data dan pengelompokkan untuk tabel dimensi serta fakta, gunakan tanda (+) pada flow kemudian pilih clean step. Hasilnya akan terlihat seperti flow dibawah ini:</p> <div></div> <p>Kemudian data di save dalam bentuk csv</p> <div><table><thead><tr><th>Pivot Tags</th><th>Tags Split</th><th>Reviews</th><th>ID_Fakta</th><th>ID_Review</th><th>Hotel_Address</th></tr></thead><tbody><tr><td>Tags - Split 6</td><td></td><td>our room was a little tight but we were comfortable the loca</td><td>1,489,001</td><td>1,489,001</td><td>14 Rue St Simon 7t</td></tr><tr><td>Tags - Split 2</td><td>'Solo traveler '</td><td>No Negative Staff</td><td>1,037,803</td><td>1,037,803</td><td>Wibautstraat 129 (</td></tr><tr><td>Tags - Split 3</td><td>'Single Room '</td><td>No Negative Staff</td><td>1,037,805</td><td>1,037,805</td><td>Wibautstraat 129 (</td></tr><tr><td>Tags - Split 4</td><td>'Stayed 2 nights ']</td><td>No Negative Staff</td><td>1,037,806</td><td>1,037,806</td><td>Wibautstraat 129 (</td></tr><tr><td>Tags - Split 5</td><td></td><td>No Negative Staff</td><td>1,037,809</td><td>1,037,809</td><td>Wibautstraat 129 (</td></tr><tr><td>Tags - Split 6</td><td></td><td>No Negative Staff</td><td>1,037,812</td><td>1,037,812</td><td>Wibautstraat 129 (</td></tr><tr><td>Tags - Split 2</td><td>'Couple '</td><td>our room was a little tight but we were comfortable the loca</td><td>1,489,000</td><td>1,489,000</td><td>14 Rue St Simon 7t</td></tr><tr><td>Tags - Split 3</td><td>'Classic Double Room '</td><td>our room was a little tight but we were comfortable the loca</td><td>1,488,999</td><td>1,488,999</td><td>14 Rue St Simon 7t</td></tr><tr><td>Tags - Split 5</td><td></td><td>our room was a little tight but we were comfortable the loca</td><td>1,488,998</td><td>1,488,998</td><td>14 Rue St Simon 7t</td></tr><tr><td>Tags - Split 4</td><td>'Stayed 3 nights ']</td><td>our room was a little tight but we were comfortable the loca</td><td>1,488,997</td><td>1,488,997</td><td>14 Rue St Simon 7t</td></tr><tr><td>Tags - Split 5</td><td>'Submitted from a mobile device ']</td><td>The pillows and bed are rough No Positive</td><td>1,952,863</td><td>1,952,863</td><td>Gran V a des les Cori</td></tr><tr><td>Tags - Split 1</td><td>['Leisure trip '</td><td>nothing Great hotel spacious room for London standards bet</td><td>1,364,544</td><td>1,364,544</td><td>2 Caxton Street We</td></tr><tr><td>Tags - Split 2</td><td>'Couple '</td><td>nothing Great hotel spacious room for London standards bet</td><td>1,364,543</td><td>1,364,543</td><td>2 Caxton Street We</td></tr><tr><td>Tags - Split 6</td><td></td><td>nothing Great hotel spacious room for London standards bet</td><td>1,364,542</td><td>1,364,542</td><td>2 Caxton Street We</td></tr><tr><td>Tags - Split 3</td><td>'Queen Room '</td><td>nothing Great hotel spacious room for London standards bet</td><td>1,364,541</td><td>1,364,541</td><td>2 Caxton Street We</td></tr><tr><td>Tags - Split 4</td><td>'Stayed 2 nights ']</td><td>nothing Great hotel spacious room for London standards bet</td><td>1,364,540</td><td>1,364,540</td><td>2 Caxton Street We</td></tr></tbody></table><div><table><thead><tr><th>Name</th><th>Date modified</th><th>Type</th><th>Size</th></tr></thead><tbody><tr><td>▼ Today</td><td></td><td></td><td></td></tr><tr><td>ProjectHotelClean.csv</td><td>25/06/2025 21:39</td><td>Excel.CSV</td><td>2.207.078 ...</td></tr><tr><td>Project</td><td>25/06/2025 21:27</td><td>Tableau Flow File</td><td>13 KB</td></tr></tbody></table></div></div>	Pivot Tags	Tags Split	Reviews	ID_Fakta	ID_Review	Hotel_Address	Tags - Split 6		our room was a little tight but we were comfortable the loca	1,489,001	1,489,001	14 Rue St Simon 7t	Tags - Split 2	'Solo traveler '	No Negative Staff	1,037,803	1,037,803	Wibautstraat 129 (Tags - Split 3	'Single Room '	No Negative Staff	1,037,805	1,037,805	Wibautstraat 129 (Tags - Split 4	'Stayed 2 nights ']	No Negative Staff	1,037,806	1,037,806	Wibautstraat 129 (Tags - Split 5		No Negative Staff	1,037,809	1,037,809	Wibautstraat 129 (Tags - Split 6		No Negative Staff	1,037,812	1,037,812	Wibautstraat 129 (Tags - Split 2	'Couple '	our room was a little tight but we were comfortable the loca	1,489,000	1,489,000	14 Rue St Simon 7t	Tags - Split 3	'Classic Double Room '	our room was a little tight but we were comfortable the loca	1,488,999	1,488,999	14 Rue St Simon 7t	Tags - Split 5		our room was a little tight but we were comfortable the loca	1,488,998	1,488,998	14 Rue St Simon 7t	Tags - Split 4	'Stayed 3 nights ']	our room was a little tight but we were comfortable the loca	1,488,997	1,488,997	14 Rue St Simon 7t	Tags - Split 5	'Submitted from a mobile device ']	The pillows and bed are rough No Positive	1,952,863	1,952,863	Gran V a des les Cori	Tags - Split 1	['Leisure trip '	nothing Great hotel spacious room for London standards bet	1,364,544	1,364,544	2 Caxton Street We	Tags - Split 2	'Couple '	nothing Great hotel spacious room for London standards bet	1,364,543	1,364,543	2 Caxton Street We	Tags - Split 6		nothing Great hotel spacious room for London standards bet	1,364,542	1,364,542	2 Caxton Street We	Tags - Split 3	'Queen Room '	nothing Great hotel spacious room for London standards bet	1,364,541	1,364,541	2 Caxton Street We	Tags - Split 4	'Stayed 2 nights ']	nothing Great hotel spacious room for London standards bet	1,364,540	1,364,540	2 Caxton Street We	Name	Date modified	Type	Size	▼ Today				ProjectHotelClean.csv	25/06/2025 21:39	Excel.CSV	2.207.078 ...	Project	25/06/2025 21:27	Tableau Flow File	13 KB
Pivot Tags	Tags Split	Reviews	ID_Fakta	ID_Review	Hotel_Address																																																																																																																		
Tags - Split 6		our room was a little tight but we were comfortable the loca	1,489,001	1,489,001	14 Rue St Simon 7t																																																																																																																		
Tags - Split 2	'Solo traveler '	No Negative Staff	1,037,803	1,037,803	Wibautstraat 129 (
Tags - Split 3	'Single Room '	No Negative Staff	1,037,805	1,037,805	Wibautstraat 129 (
Tags - Split 4	'Stayed 2 nights ']	No Negative Staff	1,037,806	1,037,806	Wibautstraat 129 (
Tags - Split 5		No Negative Staff	1,037,809	1,037,809	Wibautstraat 129 (
Tags - Split 6		No Negative Staff	1,037,812	1,037,812	Wibautstraat 129 (
Tags - Split 2	'Couple '	our room was a little tight but we were comfortable the loca	1,489,000	1,489,000	14 Rue St Simon 7t																																																																																																																		
Tags - Split 3	'Classic Double Room '	our room was a little tight but we were comfortable the loca	1,488,999	1,488,999	14 Rue St Simon 7t																																																																																																																		
Tags - Split 5		our room was a little tight but we were comfortable the loca	1,488,998	1,488,998	14 Rue St Simon 7t																																																																																																																		
Tags - Split 4	'Stayed 3 nights ']	our room was a little tight but we were comfortable the loca	1,488,997	1,488,997	14 Rue St Simon 7t																																																																																																																		
Tags - Split 5	'Submitted from a mobile device ']	The pillows and bed are rough No Positive	1,952,863	1,952,863	Gran V a des les Cori																																																																																																																		
Tags - Split 1	['Leisure trip '	nothing Great hotel spacious room for London standards bet	1,364,544	1,364,544	2 Caxton Street We																																																																																																																		
Tags - Split 2	'Couple '	nothing Great hotel spacious room for London standards bet	1,364,543	1,364,543	2 Caxton Street We																																																																																																																		
Tags - Split 6		nothing Great hotel spacious room for London standards bet	1,364,542	1,364,542	2 Caxton Street We																																																																																																																		
Tags - Split 3	'Queen Room '	nothing Great hotel spacious room for London standards bet	1,364,541	1,364,541	2 Caxton Street We																																																																																																																		
Tags - Split 4	'Stayed 2 nights ']	nothing Great hotel spacious room for London standards bet	1,364,540	1,364,540	2 Caxton Street We																																																																																																																		
Name	Date modified	Type	Size																																																																																																																				
▼ Today																																																																																																																							
ProjectHotelClean.csv	25/06/2025 21:39	Excel.CSV	2.207.078 ...																																																																																																																				
Project	25/06/2025 21:27	Tableau Flow File	13 KB																																																																																																																				
	Analisis																																																																																																																						

2.2 Proses 9 Step Kimbal

1. Memilih Proses Bisnis (Choose the Business Process)

Fokus proses bisnis adalah analisis ulasan hotel oleh pelanggan untuk meningkatkan kualitas layanan dan strategi penjualan hotel di Eropa. Manajemen hotel ingin memahami pola ulasan pelanggan berdasarkan lokasi, waktu, dan profil reviewer.

2. Menentukan Grain (Declare the Grain)

Grain ditentukan pada tingkat per ulasan individu, yang akan dianalisis berdasarkan dimensi-dimensi berikut:

- Lokasi hotel
- Pengguna/reviewer
- Waktu ulasan (tanggal, bulan, tahun, triwulan)
- Hotel

Granularity data: Setiap record adalah satu ulasan dari satu pengguna terhadap satu hotel pada satu waktu tertentu.

3. Mengidentifikasi Dimensi (Identify the Dimensions)

Dimensi yang dibutuhkan untuk menjelaskan fakta-fakta yang akan dianalisis:

Dimensi	Atribut Kunci	Deskripsi
Lokasi	Hotel_Address, Country	Lokasi hotel tempat ulasan diberikan
Reviewer	Reviewer_Nationality, User_ID	Informasi identitas dan asal pengguna
Waktu	Review_Date	Informasi waktu ulasan (tanggal, bulan, tahun, triwulan)
Hotel	Hotel_Name, Hotel_ID	Nama dan ID unik hotel yang diulas
Review	Review_Text, Reviewer_Score	Isi dan nilai skor dari ulasan yang diberikan oleh pengguna

4. Mengidentifikasi Fakta (Identify the Facts)

Fakta-fakta kuantitatif yang bisa dianalisis:

Fakta	Tipe	Sumber Perhitungan
Jumlah_Ulasan	Count	Banyaknya ulasan per kombinasi dimensi
Rata_rata_Ulasan	Numeric	Panjang rata-rata ulasan (jumlah kata/karakter)
Rata_rata_Skor_Ulasan	Numeric	Nilai rata-rata reviewer_score
Total_Ulasan	Sum	Total kata atau skor dari ulasan

5. Menyimpan Pre-Kalkulasi pada Tabel Fakta (Store Precalculated Facts)

Tabel fakta menyimpan metrik yang telah dihitung sebelumnya:

Field Fakta	Rumus SQL
CountReviews	COUNT(*) dari data ulasan
AvgReviews	AVG(LEN(Review_Text))
AvgScore	AVG(Reviewer_Score)
SumReviews	SUM(LEN(Review_Text)) atau total skor

6. Melengkapi Tabel Dimensi (Build the Dimension Tables)

Contoh struktur dan atribut dari masing-masing dimensi:

a. Dim_Lokasi

Field	Tipe Data	Deskripsi
Location_ID	INT	ID lokasi
Hotel_Address	TEXT	Alamat hotel
Country	TEXT	Negara

b. Dim_Reviewer

Field	Tipe Data	Deskripsi
Reviewer_ID	INT	ID reviewer
Reviewer_Nationality	TEXT	Kebangsaan reviewer

c. Dim_Waktu

Field	Tipe Data	Deskripsi
Date_ID	DATE	Tanggal review
Month	INT	Bulan review
Year	INT	Tahun review
Quarter	INT	Triwulan (Q1–Q4)

d. Dim_Hotel

Field	Tipe Data	Deskripsi
Hotel_ID	INT	ID hotel
Hotel_Name	TEXT	Nama hotel

e. Dim_Review

Field	Tipe Data	Deskripsi
Review_ID	INT	ID ulasan
Review_Text	TEXT	Isi teks ulasan
Reviewer_Score	FLOAT	Skor nilai ulasan dari pengguna

7. Menentukan Durasi Database (Determine the Duration of the Database)

Durasi data dalam data warehouse sangat penting untuk memastikan efektivitas dan efisiensi analisis yang dilakukan. Dalam proyek ini, durasi ditentukan berdasarkan periode waktu yang relevan dan representatif terhadap kebutuhan analisis ulasan hotel di Eropa.

Business Intelligence

Rentang Waktu Dataset

Dataset “515k Hotel Reviews in Europe” mencakup data ulasan hotel sejak tahun 2015. Namun, untuk tahap awal implementasi data warehouse, hanya data dari tahun 2015 hingga 2017 yang dimasukkan ke dalam sistem.

Tahun Data Asli	Periode Dimuat ke Data Warehouse	Durasi Tersimpan
Mulai dari 2015	2015 – 2017	2 Tahun

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Hotel Address																
2	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	08/03/2017	7.7	Hotel Arena	Russia	don't mind the broken wind	397	1403	ie park outside of the hotel was be	11	7	2.9	'Couple', 'Duplex Double Room', '0 days	523.605.759	49.159.683	
3	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	08/03/2017	7.7	Hotel Arena	Ireland	No Negative	0	1403	ie a new one Bit confusing Second	105	7	7.5	'Couple', 'Duplex Double Room', '0 days	523.605.759	49.159.683	
4	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	7/31/2017	7.7	Hotel Arena	Australia	steps So ask for single level i	42	1403	were ok It is cute hotel the breakf	21	9	7.1	'Duplex Double Room', 'Stayed 3 days	523.605.759	49.159.683	
5	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	7/31/2017	7.7	Hotel Arena	United Kingdom	insects running on the floor	210	1403	arent are nice and have a lovely out	26	1	3.8	'o traveler', 'Duplex Double Room', '3 days	523.605.759	49.159.683	
6	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	7/24/2017	7.7	Hotel Arena	New Zealand	on the photos together leave	140	1403	ng location and building Romantic	8	3	6.7	'Suite', 'Stayed 2 nights', 'Submit	10 days	523.605.759	49.159.683
7	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	7/24/2017	7.7	Hotel Arena	Poland	is total mess shouldn't happen	17	1403	reat chill out place Great park near	20	1	6.7	'Group', 'Duplex Double Room', '10 days	523.605.759	49.159.683	
8	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	7/17/2017	7.7	Hotel Arena	United Kingdom	made bed They also didn't t	33	1403	id bright The hotel is located in a q	18	6	4.6	'Twin Room', 'Stayed 5 nights', '5	17 days	523.605.759	49.159.683
9	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	7/17/2017	7.7	Hotel Arena	United Kingdom	price for the breakfast every	11	1403	ark friendly staff Food high quality	19	1	10	'Double Room', 'Stayed 2 nights', '17 days	523.605.759	49.159.683	
10	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/09/2017	7.7	Hotel Arena	Belgium	was quit dirty and outlived Al	34	1403	No Positive	0	3	6.5	'Double Room', 'Stayed 1 night', '25 days	523.605.759	49.159.683	
11	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/08/2017	7.7	Hotel Arena	Norway	akes so much noise and its ha	15	1403	a the hotel there is a big park whic	50	1	7.9	'Couple', 'Large King Room', 'Sti	26 days	523.605.759	49.159.683
12	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/07/2017	7.7	Hotel Arena	United Kingdom	Nothing all great	5	1403	new was a bit pricey but there we	101	2	10	'Double Room', 'Stayed 2 nights', '27 days	523.605.759	49.159.683	
13	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/06/2017	7.7	Hotel Arena	France	he wrong direction No late d	75	1403	Style location rooms	4	12	5.8	'plex Twin Room', 'Stayed 1 night	28 days	523.605.759	49.159.683
14	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/06/2017	7.7	Hotel Arena	United Kingdom	rooms i had a 20yr old tv in my	28	1403	Comfy bed good location	6	7	4.6	'x Double Room', 'Stayed 2 nights	28 days	523.605.759	49.159.683
15	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/04/2017	7.7	Hotel Arena	Italy	No Negative	0	1403	a large double paned glass windo	59	6	9.2	'lo traveler', 'Duplex Double Room	30 days	523.605.759	49.159.683
16	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/04/2017	7.7	Hotel Arena	Canada	itted once but that wouldn't s	35	1403	ood very historic building that s w	15	1	8.8	'n', 'Large King Room', 'Stayed 5 n	30 days	523.605.759	49.159.683
17	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/03/2017	7.7	Hotel Arena	Italy	No Negative	0	1403	easy The hotel inside is awesome	82	26	10	'Double Room', 'Stayed 1 night', '31 days	523.605.759	49.159.683	
18	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	07/03/2017	7.7	Hotel Arena	United Kingdom	ins and was put in another ide	38	1403	zing building Park location Amazin	14	8	6.3	'Duplex Double Room', 'Stayed	31 days	523.605.759	49.159.683
19	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	6/30/2017	7.7	Hotel Arena	Ireland	space a bit and the bathroom	59	1403	and easy to handle The hotel is lo	64	2	7.5	'Couple', 'Duplex Double Room', '34 days	523.605.759	49.159.683	
20	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	6/29/2017	7.7	Hotel Arena	Netherlands	No Negative	0	1403	is broken and the drains in the bat	33	4	7.1	'Group', 'Duplex Double Room', '35 days	523.605.759	49.159.683	
21	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	6/20/2017	7.7	Hotel Arena	Australia	drill The tea coffee facilities	73	1403	ed i liked their breakfast It was a q	48	16	7.5	'Double Room', 'Stayed 4 nights', '44 days	523.605.759	49.159.683	
22	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	6/19/2017	7.7	Hotel Arena	United Kingdom	ingrid not even via the tv s So	40	1403	a away Good continental breakfast	17	1	6.3	'Duplex Double Room', 'Stayed 2 n	45 days	523.605.759	49.159.683
23	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	06/12/2017	7.7	Hotel Arena	France	a smell of the gutter was very	92	1403	the whole set up with its view ar	75	12	3.8	'Duplex Double Room', 'Stayed	52 days	523.605.759	49.159.683
24	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	06/10/2017	7.7	Hotel Arena	Germany	rest to complain anymore All	364	1403	your speed The hotel locates in a l	37	21	5.4	'illy with young children', '2 rooms	54 days	523.605.759	49.159.683
25	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/29/2017	7.7	Hotel Arena	Ireland	on route to your destination t	22	1403	a and all the relevant transportatio	51	1	9.6	'Double Room', 'Stayed 2 nights', '66 days	523.605.759	49.159.683	
26	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/27/2017	7.7	Hotel Arena	United Kingdom	Nothing	3	1403	the Metro close by 8min walk or t	130	2	9.6	'Double Room', 'Stayed 4 nights', '68 days	523.605.759	49.159.683	
27	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/25/2017	7.7	Hotel Arena	ited States of America	very not enough light at the	57	1403	autiful old building former orphan	43	4	8.3	'Double Room', 'Stayed 4 nights', '70 days	523.605.759	49.159.683	
28	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/25/2017	7.7	Hotel Arena	United Kingdom	leave so if you're thinking of	51	1403	ampoo Body wash there alongside	134	2	9.6	'Group', 'Duplex Double Room', '70 days	523.605.759	49.159.683	
29	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/22/2017	7.7	Hotel Arena	France	im so it suits for business peo	31	1403	Great hotel original concept style	6	25	8.3	'olo traveler', 'Duplex Twin Room	73 days	523.605.759	49.159.683
30	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/21/2017	7.7	Hotel Arena	ited States of America	precious with no handrail	53	1403	very good and the adjacent park is	22	4	8.3	'Couple', 'Duplex Double Room', '74 days	523.605.759	49.159.683	
31	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/16/2017	7.7	Hotel Arena	Hungary	the hotel's restaurant is nic	24	1403	sture There is a big park just oppo	49	2	9.2	'Couple', 'Duplex Double Room', '79 days	523.605.759	49.159.683	
32	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	5/16/2017	7.7	Hotel Arena	Ireland	Little bit on the pricey side	7	1403	staurant area is beautiful Located i	28	5	9.2	'Double Room', 'Stayed 3 nights', '79 days	523.605.759	49.159.683	
33	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	05/11/2017	7.7	Hotel Arena	United Kingdom	ital you can still see the hosp	61	1403	staurant area is lovely and overloo	16	14	7.1	'trip', 'Group', '2 rooms', 'Stayed	84 days	523.605.759	49.159.683
34	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	05/11/2017	7.7	Hotel Arena	ited States of America	elling off Very expensive and	68	1403	No Positive	0	1	4.2	'Twin Room', 'Stayed 3 nights', '5	84 days	523.605.759	49.159.683
35	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	05/09/2017	7.7	Hotel Arena	Netherlands	Please see above	4	1403	perfect However the staff did recd	40	1	6.7	'h young children', 'Duplex Double	86 days	523.605.759	49.159.683
36	s Gravesandstraat 55 Oost 1092 AA Amsterdam Netherlands	194	05/02/2017	7.7	Hotel Arena	ited States of Amer	was basic and unwelcomin	26	1403	or two people I love the access to	32	6	7.9	'idren', '2 rooms', 'Stayed 3 night	93 days	523.605.759	49.159.683

8. Tahapan Preprocessing Data dan ETL (ETL Process and Data Cleaning)

Proses ETL dilakukan dengan Tableau Prep, terdiri dari:

a. Extract

- Mengambil data dari Kaggle
- Konversi ke format CSV atau Excel untuk diproses

b. Transform

- Menghapus data null atau tidak valid (contoh: reviewer score kosong)
- Membersihkan teks ulasan
- Membagi tanggal menjadi tahun, bulan, triwulan

Business Intelligence

- Menghitung panjang review (jumlah karakter/kata)
- Normalisasi nama negara/hotel

c. Load

- Memasukkan data bersih ke dalam tabel fakta dan dimensi
- Dapat dimuat ke database SQL Server, PostgreSQL, atau Tableau Data Extract

9. Iterasi dan Pengembangan Lanjutan (Deliver and Iterate)

Setelah data warehouse pertama selesai:

- Dashboard dapat dibuat di Power BI atau Tableau
- Analisis lanjutan: sentimen ulasan, tren waktu, preferensi negara tertentu
- Iterasi dilakukan berdasarkan feedback pengguna (manajemen hotel atau analis)

2.3 Perancangan Star Schema

BAB 3
KESIMPULAN DAN SARAN