

## ANALISIS EKSPLORASI DAN VISUALISASI PROFIL *SUPERHOST* AIRBNB KOTA MADRID DAN AMSTERDAM

Irwan Setiawan<sup>1</sup>

<sup>1</sup>Jurusan Teknik Komputer dan Informatika Politeknik Negeri Bandung  
Email: <sup>1</sup>irwan@jtk.polban.ac.id

### Abstrak

*Superhost* Airbnb merupakan *host* yang berpengalaman dan memberikan layanan yang sangat baik kepada para konsumennya. *Superhost* memiliki keistimewaan yang dapat meningkatkan jumlah pesanan dan pendapatannya. Profil *superhost* menjadi salah satu hal yang dapat dijadikan rujukan bagi *host* yang lain dalam rangka meningkatkan kualitas layanannya. Pada penelitian ini dilakukan analisis eksplorasi dan visualisasi data *host* Airbnb kota Amsterdam dan kota Madrid untuk mengetahui profil dari *superhost* berdasarkan aspek harga dan ulasan konsumen. Pemilihan kota Amsterdam dan kota Madrid dikarenakan kedua kota tersebut merupakan kota tujuan utama wisatawan di Eropa. Penelitian dilakukan dengan menggunakan pendekatan *machine learning* yang memiliki empat langkah pengerjaan, yaitu pemahaman proses bisnis, pengambilan data, pemrosesan data, dan analisis eksplorasi dan visualisasi data. *Tools* yang digunakan dalam penelitian ini adalah *Jupyter Notebook* dengan bahasa pemrograman *Python*. Hasil yang didapatkan dari penelitian ini adalah *superhost* di kota Madrid paling banyak menawarkan harga sewa pada rentang harga \$60 - \$80. *Superhost* pada rentang harga tersebut mendapatkan ulasan paling tinggi dari pelanggan berdasarkan kebersihan, komunikasi, dan *check-in*. Sedangkan untuk kota Amsterdam, *superhost* paling banyak menawarkan harga sewa pada rentang harga diatas \$140. *Superhost* pada rentang harga tersebut mendapatkan ulasan paling tinggi dari pelanggan pada semua kelompok ulasan.

**Kata Kunci:** *Airbnb, analisis, profil, superhost, visualisasi*

### Abstract

*Superhost* Airbnb is an experienced host and provides excellent service to its customers. *Superhost* has features that can increase the number of bookings and revenue. *SuperHost* profile is one thing that can be used as a reference for other hosts to improve the quality of service. In this study, an analysis of exploration and visualization of Airbnb hosts' data in the city of Amsterdam and the city of Madrid to find out the profile of *superhost* from the aspect of price and consumer reviews. The city of Amsterdam and the city of Madrid are chosen because the two cities are the leading destinations for tourists in Europe. The study was conducted using a machine learning approach that has four work steps, namely understanding business processes, data retrieval, data processing, and exploratory analysis and data visualization. The tools used in this study are *Jupyter Notebook* with the *Python* programming language. The results obtained from this study are *superhost* in Madrid, mostly offering rental prices in the price range of \$60 - \$80. They get the highest reviews from customers based on cleanliness, communication, and *check-in*. As for Amsterdam, the *superhost* offers the most rental prices in the price range above \$ 140. *Superhosts* in this price range gets the most reviews from customers in all review groups.

**Keywords:** *Airbnb, analysis, profile, superhost, visualization*

## I. PENDAHULUAN

*Superhost* merupakan status yang diberikan Airbnb kepada *host* yang berpengalaman dan memberikan layanan yang sangat baik kepada para konsumennya (Gunter, 2018). Airbnb akan mengevaluasi kinerja *Superhost* empat kali dalam setahun untuk memastikan setiap *Superhost* memberikan layanan terbaik untuk setiap pelanggannya. *Host* yang memiliki status sebagai *Superhost* memiliki beberapa keistimewaan, antara lain adalah label *Superhost* dapat dilihat dengan jelas oleh calon konsumen potensial sebagai tanda kualitas akomodasi dan komitmen dari *host*. Hal ini akan meningkatkan jumlah pemesanan dan meningkatkan pendapatan dari *host* tersebut (Liang dkk., 2017).

Keistimewaan yang dimiliki oleh *superhost* menjadi tujuan dari *host-host* lainnya, baik yang sudah lama terdaftar di Airbnb ataupun *host* baru. Profil *superhost* menjadi salah satu hal yang dapat dijadikan rujukan bagi *host* yang lain dalam rangka meningkatkan kualitas layanannya. Kota Amsterdam, Berlin, London, dan Madrid merupakan empat kota utama di benua Eropa yang menjadi tujuan wisata (Gunter, 2018).

Salah satu cara untuk mengetahui profil dari *superhost* adalah dengan menggunakan teknik analisis eksplorasi dan visualisasi data. Analisis eksplorasi dan visualisasi data telah banyak digunakan dalam berbagai penelitian untuk membantu para pengambil keputusan dalam menentukan strategi bisnis (Diamond & Mattia, 2017). Teknik ini juga digunakan pada penelitian lain seperti analisis *host* liburan singkat di Singapura (Setiawan, 2020), memetakan ketersediaan Airbnb di kota-kota di Eropa (Adamiak, 2018), memodelkan permintaan Airbnb di kota New York (Gunter dkk., 2020), dan visualisasi pengalaman tamu (Sutherland & Kiatkawsin, 2020). Analisis eksplorasi dan visualisasi data dapat dilakukan dengan menggunakan berbagai *tool*. Beberapa *tool* yang dapat digunakan untuk melakukan analisis eksplorasi dan visualisasi data adalah *python* (Kui dkk., 2017), *R* (Brownstein dkk., 2019), dan *Tableau* (Hoelscher & Mortimer, 2018).

Pada penelitian ini dilakukan analisis eksplorasi dan visualisasi data *superhost* Airbnb dari kota Amsterdam dan Madrid dari aspek harga dan ulasan konsumen dengan harapan profil *superhost* yang didapatkan dari kedua kota tersebut dapat menjadi model untuk kota-kota lainnya. *Tool* yang digunakan pada penelitian ini

adalah *Jupyter Notebook* dengan menggunakan bahasa *Python*.

## II. METODE

Rancangan penelitian terdiri dari empat tahapan. Seperti dapat dilihat pada Gambar, tahapan pertama adalah pemahaman proses bisnis. Pada tahapan ini dilakukan kajian Pustaka dan penelusuran dokumen yang berkaitan dengan proses bisnis *peer-to-peer rental*, khususnya Airbnb.

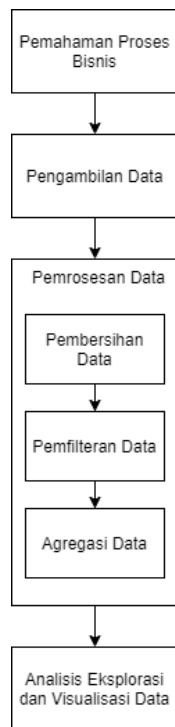
Tahapan kedua adalah pengambilan data yang berkaitan dengan *host* Airbnb. Data untuk penelitian ini diambil dari *website* *insideairbnb.com* yang menyediakan data hasil *scraping* dari *site* Airbnb. Pada penelitian ini digunakan *dataset listings\_detailed.csv* untuk kebutuhan analisis kota Madrid dan *listings\_details.csv* untuk kota Amsterdam. *Dataset* Madrid memiliki 20.837 baris dan 106 atribut dengan ukuran file 76.391 KB sedangkan *dataset* Amsterdam memiliki 20.030 baris dan 96 atribut dengan ukuran file 79.326 KB.

Tahap ketiga adalah pemrosesan data. Data yang sudah didapatkan perlu dipastikan kesiapan datanya untuk kebutuhan analisis lebih lanjut (Setiawan, 2020). Pada tahapan ini terdapat tiga sub tahapan pekerjaan, yaitu pembersihan data, filterisasi data, dan agregasi data. Pembersihan data dilakukan untuk menghilangkan anomali dan mendapatkan koleksi data yang merupakan representasi dunia nyata (Ridzuan & Wan Zainon, 2019). Pemfilteran data dilakukan untuk memfilter data yang akan digunakan dalam penelitian.

Pada penelitian ini hanya digunakan 16 atribut. Pemilihan atribut disesuaikan dengan tujuan penelitian yang lebih berfokus kepada profil *superhost* berdasarkan biaya dan ulasan dari pelanggan. Seperti dapat dilihat pada Tabel , atribut-atribut yang berkaitan dengan harga, seperti *price*, *cleaning\_fee*, *security\_deposit*, dan *extra\_people* memiliki tipe data *object* dikarenakan terdapat tanda "\$" pada setiap nilainya. Pada atribut-atribut ini dilakukan penghilangan karakter "\$" dan pengubahan tipe data menjadi *float64*.

Berkenaan dengan nilai *NULL* (*missing value*), seperti dapat dilihat pada Tabel , terdapat 10 atribut yang memiliki nilai *NULL* dengan jumlah yang beragam. Atribut *security\_deposit* dan *cleaning\_fee* memiliki jumlah *NULL* yang paling banyak. Untuk kedua atribut ini, penulis mengganti nilai *NULL* dengan angka 0 (*no*) dengan asumsi *host* tidak memintakan biaya

*security\_deposit* dan *cleaning\_fee* kepada calon konsumennya. Nilai *NULL* pada atribut *bathrooms*, *bedrooms*, *beds*, dan *host\_is\_superhost* dihilangkan dari dataset dikarenakan jumlahnya yang sangat sedikit sehingga tidak akan mengganggu hasil analisis. Untuk nilai *NULL* pada atribut *review\_scores\_location*, *review\_scores\_communication*, *review\_scores\_cleanliness* dan *review\_scores\_checkin*, tidak dilakukan aksi apapun dikarenakan dimungkinkan konsumen tidak memberikan *ulasan*.



Gambar 1. Langkah pengerjaan

Berkenaan dengan atribut *price*, penulis membuat atribut baru untuk mengelompokkan harga kedalam delapan kelompok, yaitu kelompok harga "0-20", "20-40", "40-60", "60-80", "80-100", "100-120", "120-140", dan "140+". Setelah dilakukan pemrosesan data, dataset kota Madrid terdiri dari 20.758 baris dan 17 atribut, sedangkan dataset kota Amsterdam terdiri dari 20.000 baris dan 17 atribut.

Tahap terakhir adalah analisis eksplorasi dan visualisasi data. Pada tahapan ini dilakukan analisis secara eksplorasi dan visualisasi terhadap atribut-atribut yang berkaitan dengan harga dan ulasan konsumen. Temuan-temuan yang didapatkan pada tahap ini dibahas pada bagian hasil dan pembahasan pada artikel ini. Tahapan pemrosesan data dan analisis eksplorasi dan

visualisasi data dilakukan dengan menggunakan Bahasa Python dengan tool *jupyter notebook*.

Tabel 1. Daftar Atribut Dataset

No	Nama Atribut	Type Data
1	<i>neighbourhood_cleansed</i>	<i>object</i>
2	<i>property_type</i>	<i>object</i>
3	<i>room_type</i>	<i>object</i>
4	<i>price</i>	<i>object</i>
5	<i>cleaning_fee</i>	<i>object</i>
6	<i>host_is_superhost</i>	<i>object</i>
7	<i>bathrooms</i>	<i>float64</i>
8	<i>security_deposit</i>	<i>object</i>
9	<i>extra_people</i>	<i>object</i>
10	<i>guests_included</i>	<i>int64</i>
11	<i>bedrooms</i>	<i>float64</i>
12	<i>beds</i>	<i>float64</i>
13	<i>review_scores_cleanliness</i>	<i>float64</i>
14	<i>review_scores_checkin</i>	<i>float64</i>
15	<i>review_scores_communication</i>	<i>float64</i>
16	<i>review_scores_location</i>	<i>float64</i>

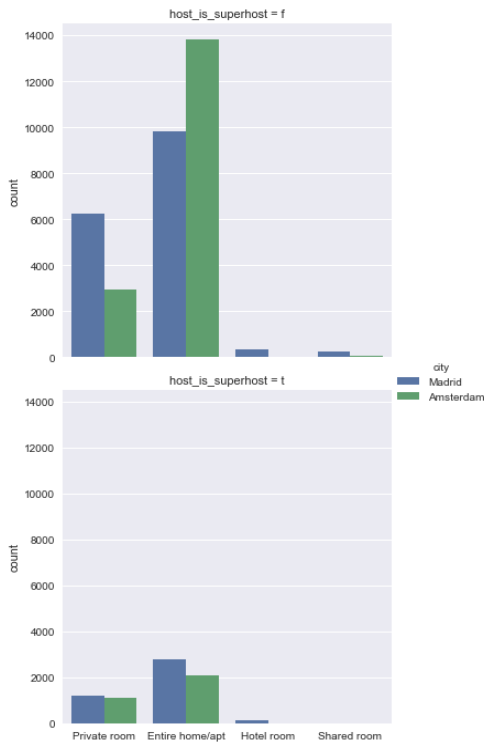
Tabel 2. Missing Value Dataset Madrid dan Amsterdam

No	Nama Atribut	Jumlah (Madrid)	Jumlah (Amsterdam)
1	<i>security_deposit</i>	6.921	6.166
2	<i>cleaning_fee</i>	5.257	3.629
3	<i>ulasan_scores_location</i>	4.968	2.661
4	<i>ulasan_scores_checkin</i>	4.965	2.660
5	<i>ulasan_scores_cleanliness</i>	4.962	2.652
6	<i>ulasan_scores_communication</i>	4.961	2.647
7	<i>beds</i>	51	10
8	<i>host_is_superhost</i>	15	8
9	<i>bathrooms</i>	10	7
10	<i>bedrooms</i>	9	4

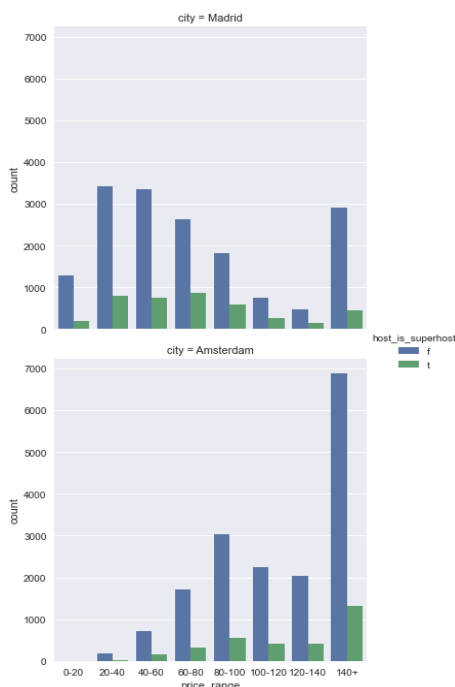
### III. HASIL DAN PEMBAHASAN

Terdapat empat tipe ruangan yang didefinisikan dalam Airbnb, yaitu *private room*, *entire home/apartment*, *hotel room*, dan *shared room*. Berdasarkan data yang ada, *entire home/apartment* merupakan tipe ruangan yang paling banyak disediakan. Jumlah *superhost* di kota Madrid lebih banyak dari kota Amsterdam, seperti ditunjukkan pada 2, dimana kota Madrid memiliki 20% *superhost* dan kota Amsterdam 16% *superhost*. Hal yang cukup menarik adalah proporsi jumlah tipe ruangan di kedua kota. Kota Madrid memiliki

proporsi *entire home/apartment* sebanyak 60,5%, *private room* sebanyak 35,9%, *hotel room* sebanyak 2,2%, dan *shared room* sebanyak 1,3%. Sedangkan kota Amsterdam hanya memiliki tiga tipe ruangan saja, yaitu *entire home/apartment* sebanyak 79,4%, *private room* sebanyak 20,3%, dan *shared room* sebanyak 0,3%.



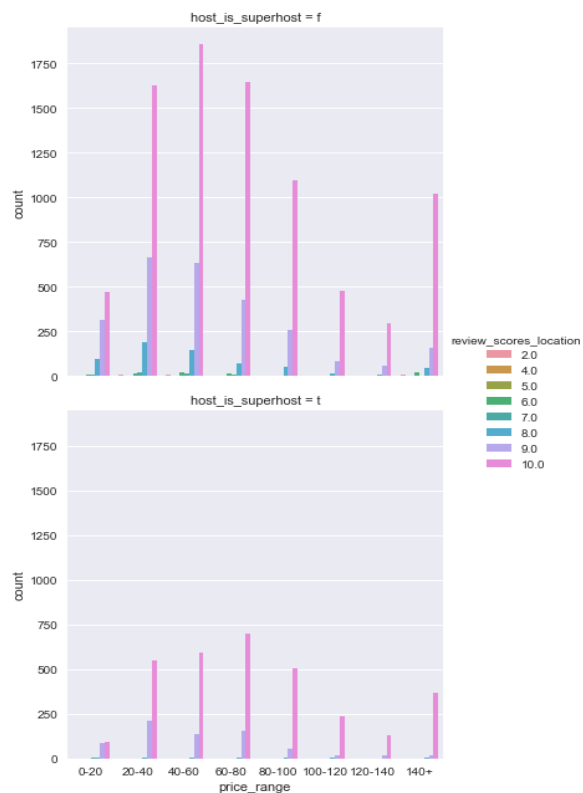
Gambar 2. Proporsi superhost berdasarkan tipe ruangan per kota



Gambar 3. Proporsi *superhost* berdasarkan harga sewa ruangan di kota Amsterdam dan Madrid

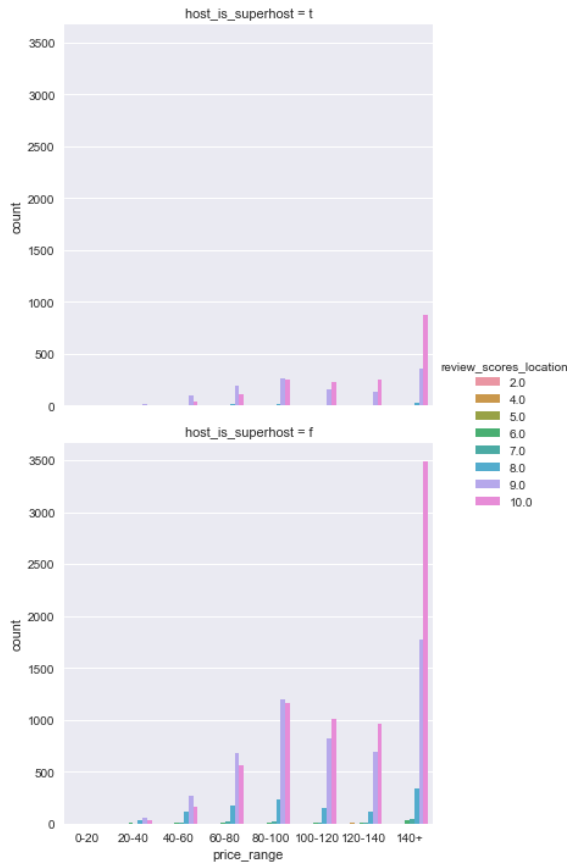
Profil *host* dan *superhost* berdasarkan aspek harga sewa ruangan, seperti dapat dilihat pada Gambar , *superhost* di kota Madrid menyewakan ruangnya pada rentang harga \$60 - \$80 dengan proporsi 21,4% pada rentang harga \$20 - \$40 dengan proporsi 19,3%, dan 18,5% pada rentang harga diatas \$60 - \$80. Sedangkan untuk kota Amsterdam, 41,4% menawarkan harga sewa ruangan pada rentang harga diatas \$140 dan 17,5% pada rentang harga \$80 - \$100. Dari Gambar , dapat dilihat bahwa *host* yang menawarkan harga sewa dibawah \$140 belum banyak yang memberikan kualitas layanan yang baik bagi para konsumennya.

Profil berdasarkan aspek ulasan terhadap lokasi yang diberikan oleh konsumen, seperti ditunjukkan pada Gambar 4, *host* yang menawarkan harga sewa pada rentang harga \$40 - \$60 mendapatkan nilai ulasan yang paling tinggi. Sedangkan *superhost* yang mendapatkan ulasan paling tinggi berada pada rentang harga \$60 - \$80.

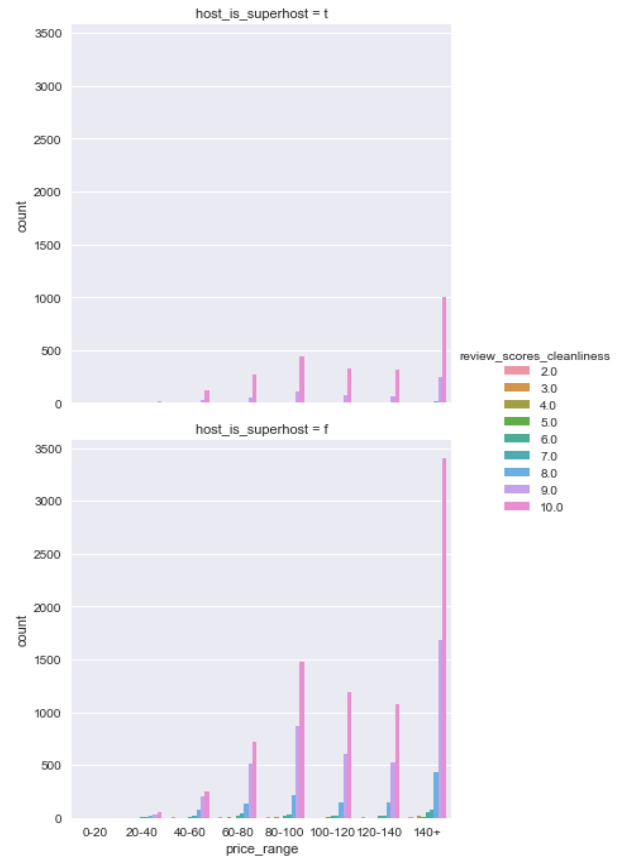


Gambar 4. Proporsi *host* dan *superhost* berdasarkan ulasan lokasi di kota Madrid

Untuk kota Amsterdam, ditunjukkan pada Gambar 5, *host* dan *superhost* pada rentang harga diatas \$140 mendapatkan ulasan berdasarkan lokasi paling tinggi. Hal ini menunjukkan bahwa lokasi *host* pada rentang harga dibawah \$140 kurang sesuai dengan harapan konsumen.



Gambar 5. Proporsi *host* dan *superhost* berdasarkan ulasan lokasi di kota Amsterdam

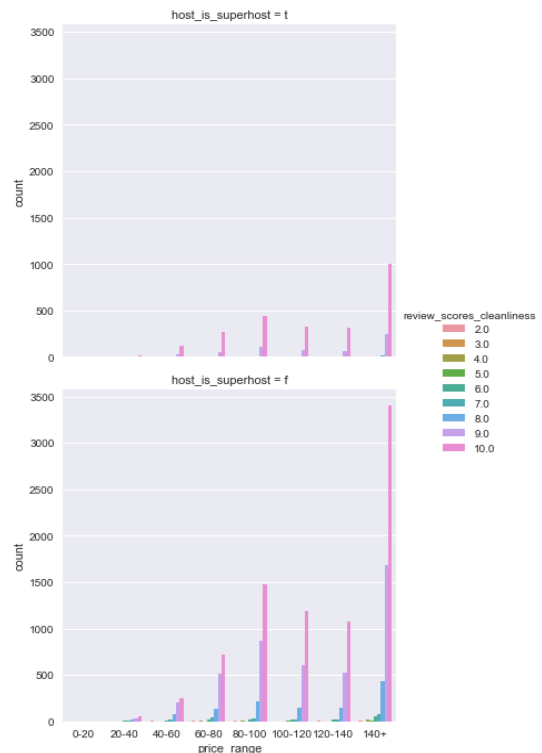


Gambar 6. Proporsi *host* dan *superhost* berdasarkan ulasan kebersihan ruangan di kota Amsterdam

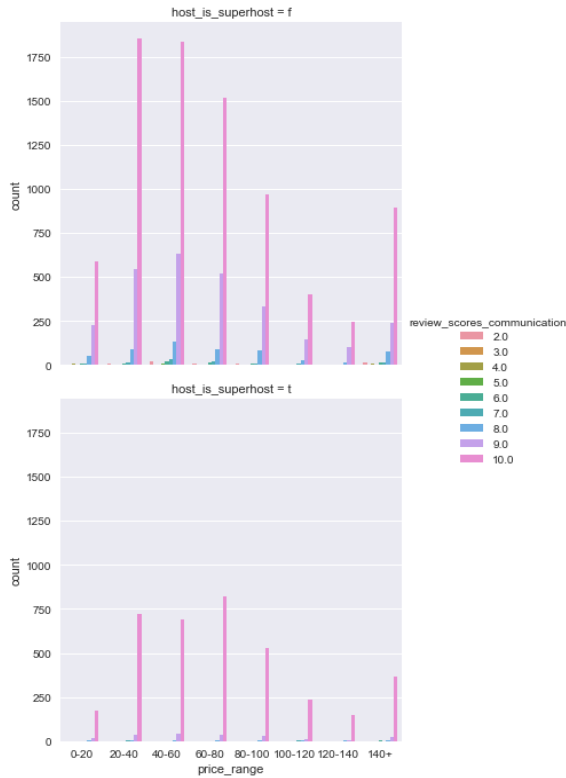
Profil berdasarkan aspek ulasan terhadap kebersihan ruangan yang diberikan oleh konsumen, seperti ditunjukkan pada Gambar 6, *host* yang menawarkan harga sewa pada rentang harga \$20 - \$40 mendapatkan nilai ulasan yang paling tinggi. Sedangkan *superhost* yang mendapatkan ulasan paling tinggi berada pada rentang harga \$60 - \$80.

Untuk kota Amsterdam, ditunjukkan pada Gambar 7, *host* dan *superhost* pada rentang harga diatas \$140 mendapatkan ulasan berdasarkan kebersihan ruangan paling tinggi. Hal ini menunjukkan bahwa kebersihan ruangan *host* pada rentang harga dibawah \$140 kurang sesuai dengan harapan konsumen.

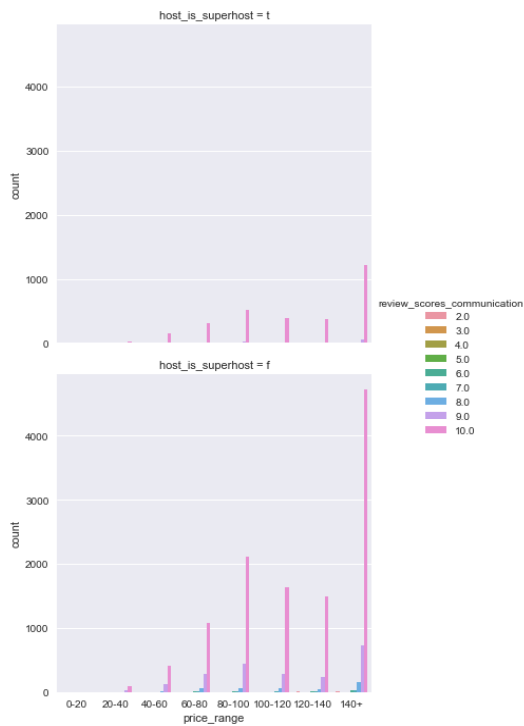
Profil berdasarkan aspek ulasan konsumen terhadap komunikasi pemilik ruangan, seperti ditunjukkan pada Gambar 8, *host* yang menawarkan harga sewa pada rentang harga \$20 - \$40 mendapatkan nilai ulasan yang paling tinggi. Sedangkan *superhost* yang mendapatkan ulasan paling tinggi berada pada rentang harga \$60 - \$80. Untuk kota Amsterdam, ditunjukkan pada Gambar, *host* dan *superhost* pada rentang harga diatas \$140 mendapatkan ulasan berdasarkan komunikasi pemilik ruangan paling tinggi.



Gambar 7. Proporsi *host* dan *superhost* berdasarkan ulasan kebersihan ruangan di kota Amsterdam



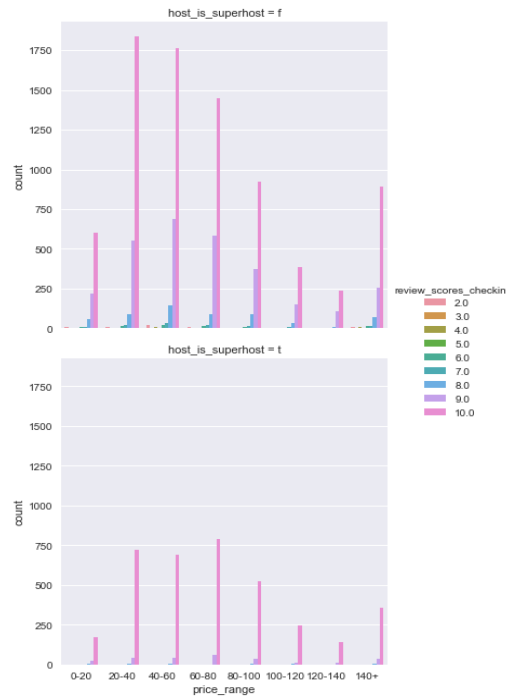
Gambar 8. Proporsi *host* dan *superhost* berdasarkan ulasan komunikasi di kota Madrid



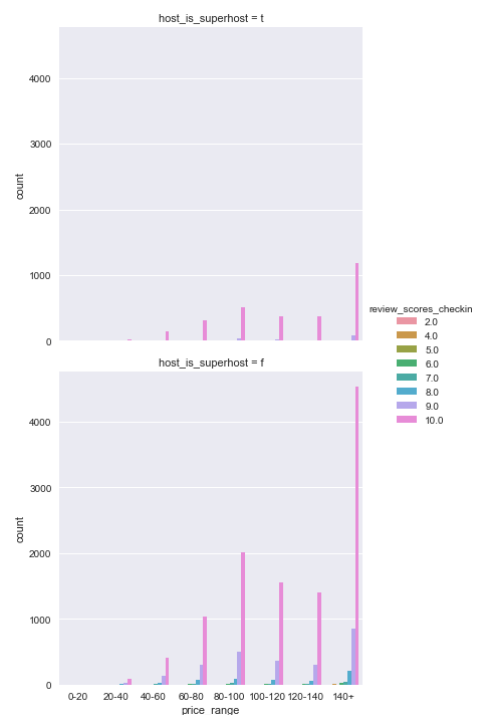
Gambar 9. Proporsi *host* dan *superhost* berdasarkan ulasan komunikasi di kota Amsterdam

Profil berdasarkan aspek ulasan terhadap pengalaman *check-in* yang diberikan oleh konsumen, seperti ditunjukkan pada Gambar 10, *host* yang menawarkan harga sewa pada rentang harga \$20 - \$40 mendapatkan nilai ulasan yang

paling tinggi. Sedangkan *superhost* yang mendapatkan ulasan paling tinggi berada pada rentang harga \$60 - \$80. Untuk kota Amsterdam, ditunjukkan pada Gambar 11, *host* dan *superhost* pada rentang harga diatas \$140 mendapatkan ulasan berdasarkan pengalaman *check-in* paling tinggi



Gambar 10. Proporsi *host* dan *superhost* berdasarkan ulasan *check-in* di kota Madrid



Gambar 11. Proporsi *host* dan *superhost* berdasarkan ulasan *check-in* di kota Amsterdam

#### IV. PENUTUP

##### Kesimpulan

Pada penelitian telah dilakukan analisis eksplorasi dan visualisasi data host Airbnb di kota Amsterdam dan kota Madrid untuk mendapatkan profil *superhost* berdasarkan aspek harga sewa dan ulasan pelanggan. Berdasarkan temuan yang didapatkan, *superhost* di kota Madrid paling banyak menawarkan harga sewa pada rentang harga \$60 - \$80. *Superhost* yang mendapatkan nilai ulasan dari pelanggan dengan nilai tertinggi paling banyak didapatkan oleh *superhost* yang menawarkan harga sewa pada rentang harga \$60 - \$80.

Untuk kota Amsterdam, *superhost* paling banyak adalah *host* yang menawarkan harga sewa pada rentang harga diatas \$140. *Superhost* pada rentang harga tersebut juga mendapatkan nilai ulasan dari pelanggan dengan nilai tertinggi yang paling banyak.

##### Saran

Penelitian ini baru mengeksplorasi profil *superhost* berdasarkan aspek harga sewa dan ulasan pelanggan saja. Agar mendapatkan profil *superhost* yang lebih lengkap, ada baiknya dilihat juga profil *superhost* berdasarkan aspek lainnya seperti deskripsi ulasan dari pelanggan, jumlah konsumen yang menginap, dan berapa lama konsumen tersebut menginap.

##### Ucapan Terima Kasih

Penelitian ini dibiayai oleh DIPA Politeknik Negeri Bandung sesuai dengan Surat Perjanjian Pelaksanaan Penelitian skema Penelitian Pemula (PP) Nomor: B/186.19/PL1.R7/PG.00.03/2020.

#### V. DAFTAR PUSTAKA

- Adamiak, C. (2018). Mapping Airbnb supply in European cities. *Annals of Tourism Research*, 71, 67–71. <https://doi.org/10.1016/j.annals.2018.02.008>
- Brownstein, N. C., Adolfsson, A., & Ackerman, M. (2019). Descriptive statistics and visualization of data from the R datasets package with implications for clusterability. *Data in Brief*, 25, 104004. <https://doi.org/10.1016/j.dib.2019.104004>
- Diamond, M., & Mattia, A. (2017). Data Visualization: An Exploratory Study into the Software Tools Used by Businesses. *Journal of Instructional Pedagogies*, 18. <https://eric.ed.gov/?id=EJ1151731>
- Gunter, U. (2018). What makes an Airbnb host a superhost? Empirical evidence from San Francisco and the Bay Area. *Tourism Management*, 66, 26–37. <https://doi.org/10.1016/j.tourman.2017.11.003>
- Gunter, U., Önder, I., & Zekan, B. (2020). Modeling Airbnb demand to New York City while employing spatial panel data at the listing level. *Tourism Management*, 77, 104000. <https://doi.org/10.1016/j.tourman.2019.104000>
- Hoelscher, J., & Mortimer, A. (2018). Using Tableau to visualize data and drive decision-making. *Journal of Accounting Education*, 44, 49–59. <https://doi.org/10.1016/j.jaccedu.2018.05.002>
- Kui, X., Liu, W., Xia, J., & Du, H. (2017). Research on the improvement of python language programming course teaching methods based on visualization. *2017 12th International Conference on Computer Science and Education (ICCSE)*, 639–644. <https://doi.org/10.1109/ICCSE.2017.8085571>
- Liang, S., Schuckert, M., Law, R., & Chen, C.-C. (2017). Be a “Superhost”: The importance of badge systems for peer-to-peer rental accommodations. *Tourism Management*, 60, 454–465. <https://doi.org/10.1016/j.tourman.2017.01.007>
- Ridzuan, F., & Wan Zainon, W. M. N. (2019). A Ulasan on Data Cleansing Methods for Big Data. *Procedia Computer Science*, 161, 731–738. <https://doi.org/10.1016/j.procs.2019.11.177>
- Setiawan, I. (2020). ANALISIS DAN VISUALISASI HOST LIBURAN SINGKAT MENGGUNAKAN PENDEKATAN ANALISIS GEOSPASIAL. *JTT (Jurnal Teknologi Terapan)*, 6(1), 1–7. <https://doi.org/10.31884/jtt.v6i1.249>
- Sutherland, I., & Kiatkawsin, K. (2020). Determinants of Guest Experience in Airbnb: A Topic Modeling Approach Using LDA. *Sustainability*, 12(8), 3402. <https://doi.org/10.3390/su12083402>