

CONFERENCE OF EUROPEAN STATISTICIANS

Expert Meeting on the Dissemination and Communication of Statistics

13-15 September 2022, Online

Using Google Maps Review Data For BPS Office Evaluation

Satria Bagus Panuntun, Wiwin Srimulyani, Maulana Faris, Setia Pramana

Statistics Indonesia; Politeknik Statistika STIS, Jakarta

satria.bagus@bps.go.id; wiwin.srimulyani@bps.go.id; maulana.faris@bps.go.id;
setia.pramana@stis.ac.id

Abstract: Every National Statistical Office (NSO) needs to evaluate the public services that have been provided to the community, including BPS as NSO Indonesia. Google currently provides facilities for users of public facilities to review and give ratings to a place, including the services of the BPS office. Big Data reviews from the public attached to the google map can be used to evaluate the performance of BPS throughout Indonesia, both local, provincial, and central BPS. This study collects rating and review data from 274 district/city BPS offices in 34 provinces in Indonesia with 2204 reviews containing comments and ratings. From this data, we can classify local BPS offices with the lowest rating to the perfect rating from users who do reviews. We can also sort the BPS city with the most reviews to the lowest. In this study, we also make a sentiment analysis of the BPS offices of districts and cities throughout Indonesia.

Keywords: *Big Data, Google Maps, Google Review*

topics: S2 Market Research and Digital Marketing and Communication Strategies

Introduction

BPS is the national statistical office in Indonesia which job is to provide data for the purposes of decision-making for both the government and the wider community. BPS is a vertical agency that has its head office in the capital city of Jakarta and has provincial headquarters in every provincial capital throughout Indonesia and almost all regencies and cities have local statistical offices. This will make it easier for BPS to serve data users in districts

and cities throughout Indonesia. However, the number of BPS offices makes the quality of their services vary from one another. BPS requires an evaluation tool to see the performance of public services from each BPS office throughout Indonesia. The internet is one of the potentials where BPS can capture public opinion on the performance of each BPS public service.

The internet is now a place where the flow of information and data is centered. Information exchange on the internet occurs every second from various places around the world. Google is a multinational company that specializes in providing internet services and products. Google has a web mapping platform and an application called google maps where each user can provide a rating and review of public places including the BPS office through a feature called Google review. Google maps itself is a web mapping service that provides satellite imagery, street maps, 360° panoramas, traffic conditions, and route planning for traveling by foot, car, bicycle (beta version), or public transportation. Google maps for mobile was released in September 2008 and went on to become one of the most popular apps in the world. Each user can easily provide a rating and review of a place via their smart phone.

In previous research, the Google review feature was used to evaluate tourism places through an assessment of visitor satisfaction. According to Haq and Rachmat (2020), in the field of tourism, Google reviews are an important aspect in the era of big data because there is a lot of trace information from tourists in one place. The location feature of the phone provides automatic notifications in the mobile phone asking the user to review a place he has just visited. The results of reviews from Google Maps users are then collected into big data. The available data and information then becomes a determinant in making a decision whether the place is good to visit or not. This review is commonly referred to as electronic-word of mouth (eWom). However, there is still no research on the use of the Google Review feature for government offices, especially national statistical offices.

BPS sees this as a potential to capture public opinion about services in local statistical offices throughout Indonesia. The ratings and reviews obtained will be used as evaluation materials to improve public services in the future.

This research aims to see:

1. What can the National Statistical Office use from the information provided by google maps for evaluating public services to data users?
2. How is the map of the distribution of local BPS offices with the best to the worst services?

Methodology

Data Collection Method

Data collection was carried out by collecting visitor review data from the BPS office work unit in Indonesia from the Google Maps website. This site was chosen as the data source because it is ranked first as top website ranking for Maps in Indonesia and it provides self-contribution

visitor's review [1]. The workflow of the data collection process from the Google Maps website can be explained in 4 steps, namely Preparation, Determination of Web Scraping Technology, User Reviews Data Collecting, and Cleaning/Compiling Data.

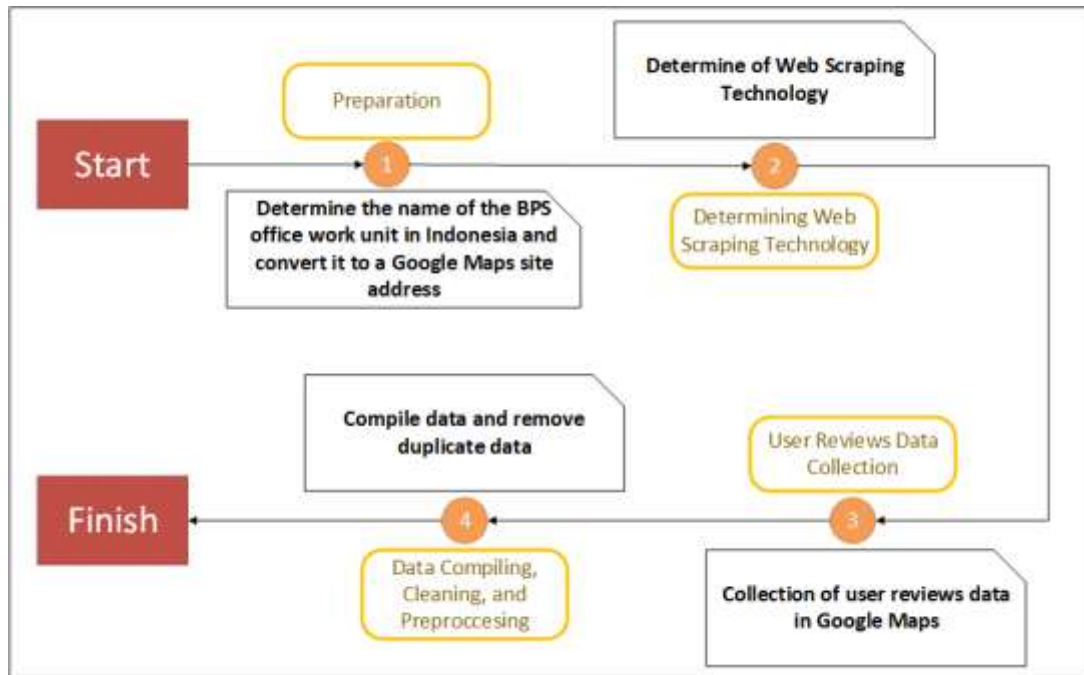


Figure 1. Workflow of the Data Collection Process from Google Maps

1. Preparation

In this study we found that on the Google Maps site, users can leave reviews of the places they visited. Based on this, we want to see if the work units of the BPS office in Indonesia have also been reviewed by the visitors. In this first stage, namely the preparation of data collection, we need to determine the name of the work unit from each BPS office at the district/city level and convert it into a search URL structure on the Google Maps site. We determine 514 district/city BPS offices in Indonesia to collect visitor review data, with the hope that we can collect all visitor reviews from each of these BPS offices. Furthermore, each district/city BPS office name will be converted into a search URL on the Google Maps site by adding "https://www.google.com/maps/search/" before the district/city BPS office name. Examples of the conversion results obtained are as follows.

District/city BPS Work Unit	Search URL Conversion Results
BPS Provinsi Aceh	https://www.google.com/maps/search/BPS+Provinsi+Aceh
BPS Kota Bandung	https://www.google.com/maps/search/BPS+Kota+Bandung
BPS Kabupaten Magelang	https://www.google.com/maps/search/BPS+Kabupaten+Magelang

2. Determination of Web Scraping Technology

After determining which BPS office work unit for which visitor review data will be collected, in this step we determine what web scraping technique will be used to collect visitor reviews data from Google Maps. We found that there are two methods to collect visitor review data on the Google Maps site. The first method is to use the API provided directly by Google and the second method is the Web Scraping method, which is to extract a review page of a place on the Google Maps site from its HTML structure. In this study, we used the second method in collecting data by doing Web Scraping from pages on the Google Maps site.

We use Python programming language with Selenium module to access every page of Google Maps site with web driver [2]. WebDriver drives a browser natively, as a user would, either locally or on a remote machine using the Selenium server, marks a leap forward in terms of browser automation [3]. By using this technique, the selenium web driver will access every search URL on the Google Maps site, as if each page was accessed by real internet users. The selenium web driver will display visitor reviews from each BPS office work unit which will be collected at a later stage. The following is an example of a selenium web driver accessing a Google Maps page.

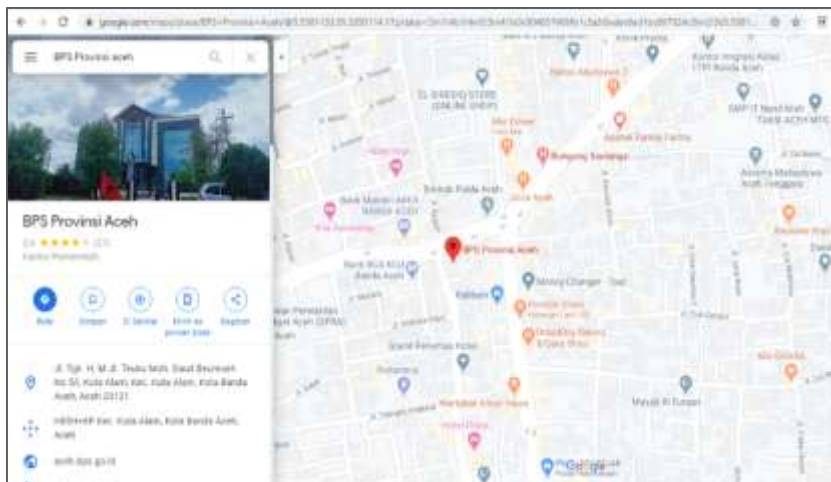


Figure 2. Selenium Web Driver Accessing a Google Maps Page

3. User Reviews Data Collecting

After the browser displays the BPS office search page on Google Maps, then the selenium web driver= will scroll the page down until it finds the "See All Reviews" button. This process can be set in python syntax with the command `driver.execute_script("window.scrollTo(0, Y)")`. When the "See All Reviews" button is clicked, a review page will appear as shown below.

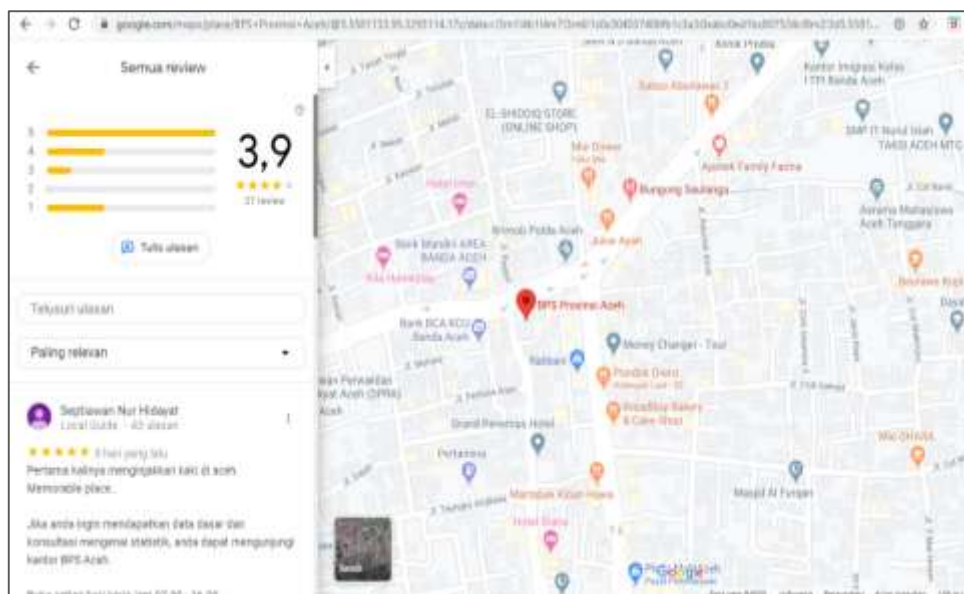


Figure 3. Review Page on Google Maps Page

The command to scroll the page down will be carried out until all visitor reviews on the page have been successfully loaded. After all the reviews are loaded, then we take the information displayed on the page. Information collected from this page includes:

- Place name
- Place address
- Visitor reviews
- Visitor ratings
- Review date

All the information above will be collected by marking the position of the HTML structure tag using the Selenium framework in Python. For example, the position of the visitor review is located at `` tag in the HTML CSS structure of this page.

4. Data Compiling, Cleaning, and Preprocessing

After doing web scraping, the collected information is then tabulated with the Pandas module. Tabulated data is then stored and compiled into a filebase by going through the cleaning process first. The cleaning process includes:

- Changing the data field into the correct data type. Changing the data field is necessary so we can analyze the result in the right way.
- Removing duplication of data based on property advertisement URL address.

The clean Google Maps data then needs to be preprocessed to get a data structure that fits the needs of the analysis method. The preprocessing stage includes:

- Translating visitor reviews in Indonesian into English using the Googletrans package in Python. The translation is done to get the sentiment of visitors based on their reviews. Sentiment analysis will be explained later.
- Standardize the name structure of the BPS office work unit by using the Named Entity Recognition (NER) method. NER is a part of Natural Language Processing (NLP) research that is used to extract information such as the name of a person, organization, location, and time from text-formatted data [4]. We use NER to reduce the name of the district/city

and the name of the province where the office is located. The structure of the office name is "BPS (city/district) (name of city/district)". For example, the name of the BPS office on Google Maps is "The Central Bureau of Statistics of Bogor", the name of this entity is then standardized to "BPS Kabupaten Bogor".

Data Analysis Method

In this study, the analysis of visitor reviews data from Google Maps is carried out in several stages. Data analysis begins with data visualization, data descriptive analysis, and sentiment analysis.

- Data Visualization

Data that has been processed will be visualized to make it more representative and interactive. The tool used to visualize the data is a desktop application, namely Microsoft PowerBI. With this application we can create various kinds of visualizations such as line charts, bar charts, pie charts, and even world map-shaped charts.

- Data Descriptive Analysis

The data set that has been tabulated and visualized is then analyzed by assessing its characteristics. These characteristics include the number of reviews from each BPS office work unit, the average rating given by visitors, and others.

- Sentiment Analysis

In addition to descriptive analysis, we also perform a sentiment classification of visitor reviews on Google Maps. Sentiment classification in this study was carried out using the Lexicon-based method using the VADER Sentiment dictionary. Sentiment analysis is a part of Natural Language Processing (NLP) research that is used to identify and extract opinions in certain texts [5]. Sentiment classification with this Lexicon-based method uses a dictionary or lexicon dictionary to evaluate the words in a text or in other words, each word in a sentence will be given a positive or negative score [6]. With this analysis, we want to know the sentiments of visitors to the BPS office work units they have visited.

I. Results and Findings

Data Collection Result

The data collection method yielded a data set with 2,442 rows of visitor review data. From this dataset, we find that only 287 district/city BPS offices have Google Maps review data, out of 514 districts/cities in Indonesia. The rest of the BPS office work units whose data were not collected may occur because there has been no Google Maps review left by visitors to the work unit. Or it may happen because the BPS office work unit is not found or has not been registered on Google Maps. Table 2 below shows the number of visitor reviews by type of review.

Review Type	Total
Reviews with comments and rating	1,146
Reviews with rating only	1,296

From table 2, it can be indicated that the majority of visitors prefer to give reviews with rating only rather than having to add comments to their reviews. If we observe based on the province where the BPS office work unit is located (Figure 4), it can be seen that the work unit that is most reviewed by visitors is the BPS office located in Java such as Central Java, East Java, and West Java with a percentage of more than 10% of the total Google Maps review data collected. This indicates that people in Java are more likely to give Google Maps reviews than people in other areas.

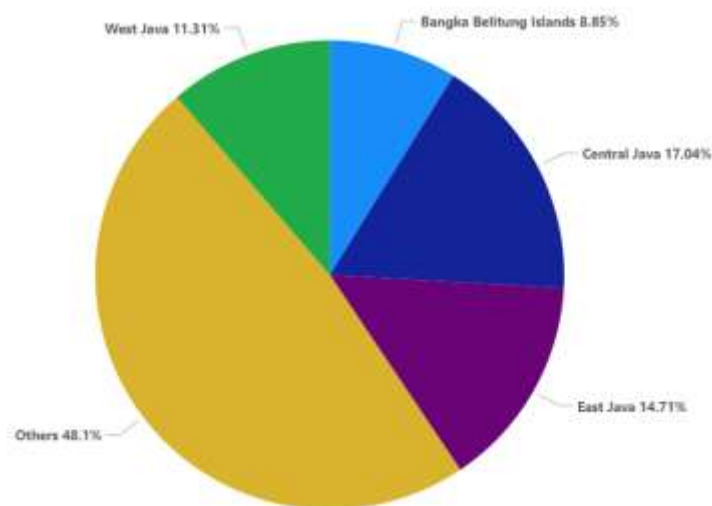


Figure 4. Number of Visitor Reviews per Province

Furthermore, if we look at the distribution of the number of visitor reviews using map visualization (Figure 5), it can be seen that not all BPS Offices either at province or district level could be collected during scraping process.

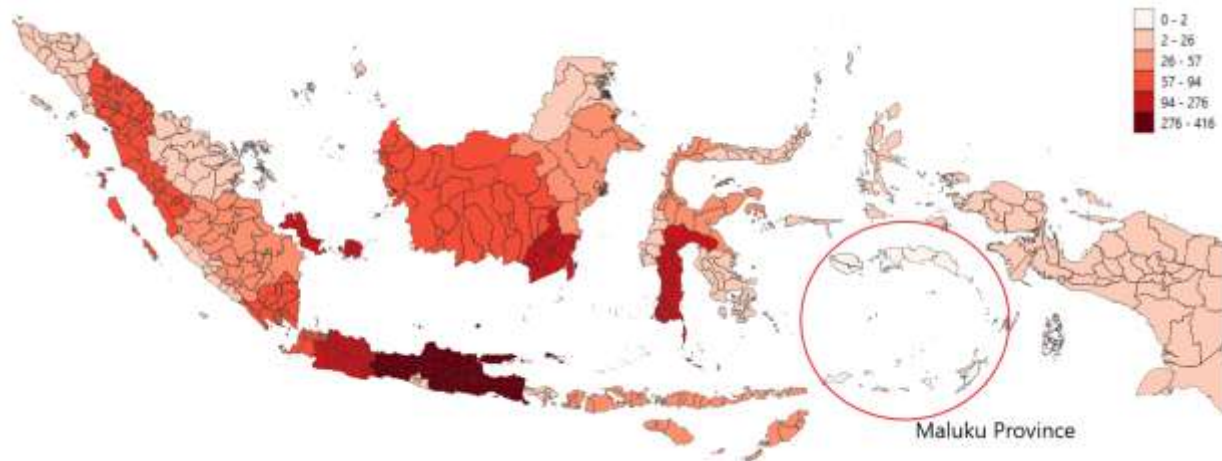


Figure 5. Number of Visitor Reviews per Province in Map

Data Analysis Results

From 2,442 Google Maps review data collected, the average rating given by visitors at 287 BPS offices in Indonesia is 4.46 out of 5. This rating is certainly high, so it can be said that the average visitor is quite satisfied with their visit to the BPS work unit in Indonesia. There are 71 BPS office work units that have a perfect rating (5 out of 5), including BPS Belitung Regency in Bangka Belitung Islands Province, BPS Asahan Regency in North Sumatra Province, and BPS Palu City in Central Sulawesi Province. While the BPS office with the lowest rating is BPS Banyu Asin Regency in South Sumatra Province with a score of 2.0 out of 5.0. Figure 6 below shows the distribution of ratings given by visitors using map visualization.

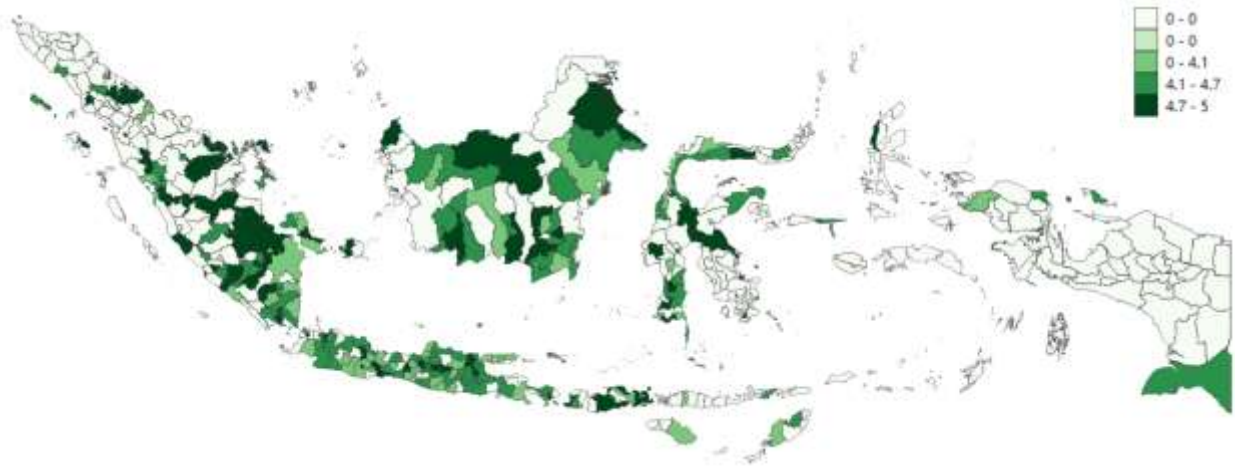


Figure 6. Distribution of Ratings Given by Visitors

We also analyzed the sentiments of comments attached to visitor reviews (Figure 7). From 1,146 visitor comments, 653 comments (56.98%) had positive sentiments and only 36 comments (3.14%) had negative sentiments, while the rest had neutral sentiments. This also further shows that the majority of visitors are satisfied with their visit to the BPS office work unit in Indonesia.

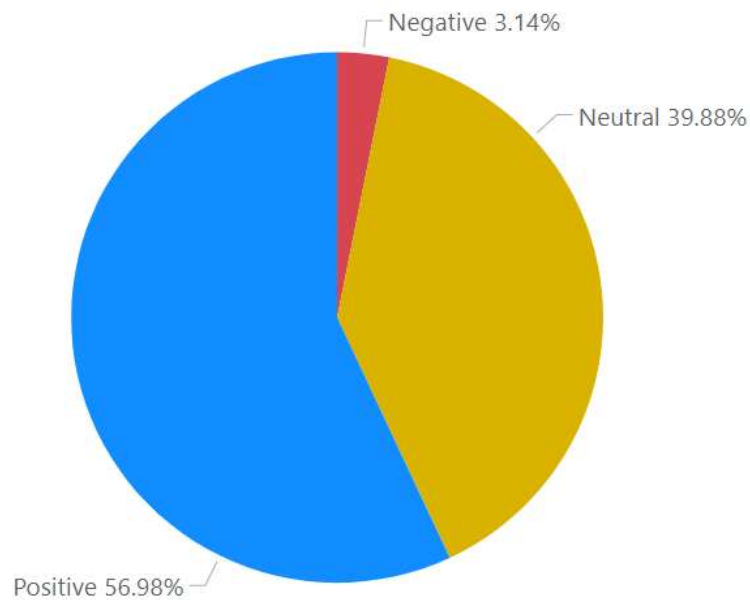


Figure 7. Visitor Review Sentiment Percentage

This sentiment analysis can also be grouped according to the province where the BPS office work unit is located. In Figure 8, there are 7 provinces whose district/city BPS offices have

negative sentiments. Among them are the Provinces of North Kalimantan, West Nusa Tenggara, Aceh, South Sumatra, East Kalimantan, West Java, and Central Java, with the highest number of negative sentiments in Central Java Province.

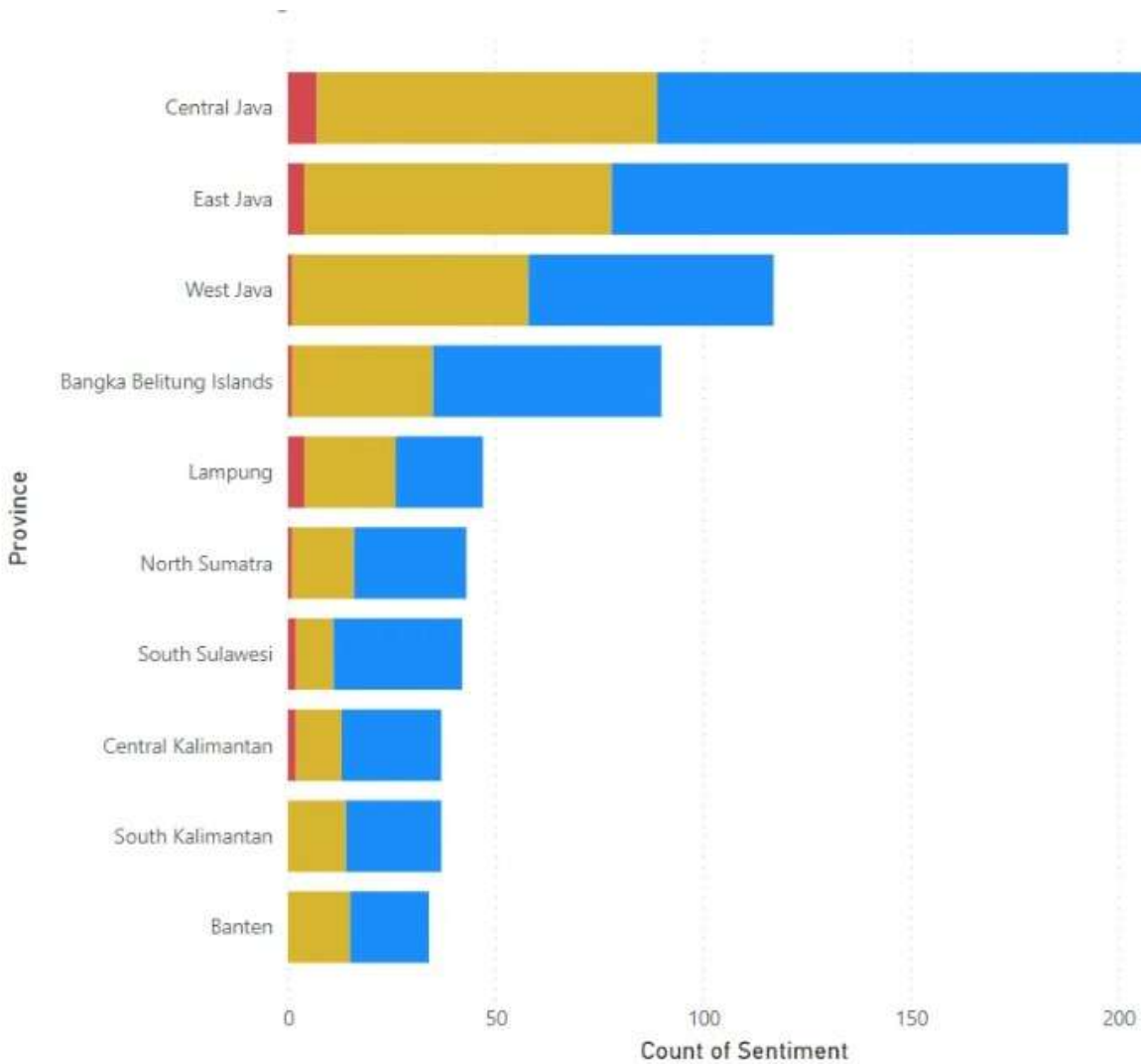


Figure 8. Visitor Review Sentiment per Province

II. Discussion

From the results of collecting data on visitor reviews of BPS office work units in Indonesia, the results are expected to be used by policy makers at the Regency/City BPS offices as an evaluation material for services and facilities in their offices. For example, if we look at some

comments from the sentiment analysis in table 3, we can find out what reviews have positive and negative sentiments from visitors.

Sentiment	Review
Positive	Comfortable and friendly
Positive	The service is good, friendly and fast
Positive	The office is clean, nice...
Positive	The Central Statistics Agency of Kudus is indeed great, the office is modern, clean, neat and beautiful...
Positive	Strategic location, service is ok
Positive	Friendly service, clean and tidy place.
Negative	Not clearly visible from the road
Negative	The place is not strategic and its position is at the end of the road, it's a shame.
Negative	It's very helpful to know areas that have never been visited. just a shame. Location/Borders of the administration area are all wrong. Example: The area where I enter Rancepato Village, but on Google Maps it is claimed to be Rancajaya Village. that's a big mistake. Please fix it...!!

It can be seen from the examples of the reviews above that visitors not only commented on the statistical data produced by BPS, but also commented on the cleanliness of the office, the location of the office on the road, and the behavior of office employees. Therefore, it can be clearly concluded that the collected Google Maps review data can be used by policy makers as evaluation material for their BPS office work units.

Unfortunately, from the data collection and analysis that has been carried out, it turns out that the results have not been able to cover all BPS office work units in Indonesia (only 287 out of 514 BPS offices). As previously explained this may happen because there has been no Google Maps review left by visitors to the work unit. Or it may happen because the BPS office work unit is not found or has not been registered on Google Maps. To overcome this, policy making at the Regency/City BPS offices is expected to register their work units on Google Maps with a standardized name so that the process of collecting visitor review data can be collected more completely. In addition, we can also use data collection methods with APIs that have been provided by Google by paying a certain amount of money. The results with this method will certainly produce more complete and accurate data.

The results of the rating and sentiment analysis carried out also cannot be confirmed, considering that anyone can leave a review on Google Maps, including employees from the district/city office itself. This of course can lead to bias from the analysis carried out.

III. Conclusion

There are a total of 2,442 reviews provided by Google Maps users who visited 287 district/city BPS offices in 33 provinces in Indonesia (Maluku Province not covered). The average rating given by visitors is 4.46 out of 5, with 653 comments with positive sentiment or about 56.98% of the total comments given by visitors. There are 71 BPS office work units that have a perfect rating (5 out of 5), with BPS office with the lowest rating is BPS Banyu Asin Regency in South Sumatra Province with a score of 2.0 out of 5.0. The results of data collection can clearly be used as evaluation material by policy makers at the district/city BPS offices because visitor comments contain not only statistical data produced by BPS, but also services and facilities at the BPS office.

REFERENCES

- Fatima, S. Luqmaan, S. Rasheed, N.A. 2021. Web Scraping with Python and Selenium. IOSR Journal of Computer Engineering (IOSR-JCE).
- Haq, Fathia U. & Rachmat, Heryadi. (2020). Penggunaan Google Review Sebagai Penilaian Kepuasan Pengunjung dalam Pariwisata. Tornare-Journal of Sustainable Tourism Research.
- Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.
- Kundi, F. M., Khan, A., Ahmad, S., & Asghar, M. Z. (2014). Lexicon-based sentiment analysis in the social web. *Journal of Basic and Applied Scientific Research*, 4(6), 238-48.
- Sahrul & Meredita Susanty, M.Sc. 2020. MODEL NAMED ENTITY RECOGNITION (NER) DALAM BAHASA INDONESIA DENGAN MENGGUNAKAN BiLSTM-CNNs Selenium Project, "Selenium Documentation," pp. 1–158, 2010, [Online]. Available: http://oss.infoscience.co.jp/seleniumhq/docs/book/Selenium_Documentation.pdf
- similarweb, "Top Websites Ranking for Maps in the world," 2022. <https://www.similarweb.com/top-websites/category/reference-materials/maps/> (accessed Sep. 05, 2022).