Abstract—Computational linguistics is concerned with understanding language from a computational perspective and constructing artifacts that are useful in processing and generating language. In the use of language, whether human language or programming language, there can be an error that makes the language not understood properly. One of the errors that often occurs is syntax error. In language, a syntax error is a mistake in using a language that involves organizing words and phrases that do not make sense. While in programming, a syntax error is an error in writing code in a program that makes the format or information unrecognizable by the computer system. Such errors are the simplest of errors but can affect many aspects of the final code output. This article aims to show how writing errors or typos in programming code can affect some or all of the results. The data obtained is data from web programming code that is used to make the website display the Geographic Information System Clustering the Distribution of Stunting Disease in Banggai Regency with K-Means, and also the R programming code used to calculate the stunting distribution using K-Means. The results of this article will show that a typo, even just a letter or any single punctuation, can affect the program’s final result.

Index Terms—computational linguistics, syntax error, typo, programming code

I. INTRODUCTION

Linguistics is a scientific study of language and its structure including the study of grammar, syntax, and phonetics. Specific branches of linguistics include sociolinguistics, dialectology, psycholinguistics, computational linguistics, comparative linguistics, and structural linguistics (Soanes & Stevenson, 2003). Computational linguistics is an idea by Chomsky (1995) which aims to install and initialize languages to produce certain programs. Computational linguistics is the scientific and engineering discipline from a computational perspective and is concerned with understanding written and spoken language from a computational perspective, and building artifacts that usefully process and produce language, either in bulk or in a dialogue setting (Schubert, 2014). Effective communication including dialogue of human life has made the language context contribute to the meaning’s interpretation (Yulianti et al., 2022). Computational linguistics also explores how human language might be automatically processed and interpreted. The human language includes the speech act of daily life by a speaker in communication, such as apologizing, complimenting, and complaining that occur anywhere at home, at workplaces, in organizations, or even on social media (Arafah & Kaharuddin, 2019). Other than that, the human language relates to the language form and its use as well as the cultural environment that is created based on how people interact and socialize where the language contains meaningful and complete information to explore before processing and interpreting a new program (Arafah et al., 2020; Fadillah et al., 2022). A broader perspective on human language in writing a literary work can also capture its cultural environment (Siwi et al., 2022). Studying literature needs interdisciplinary, multidisciplinary, and trans-disciplinary perspectives, which can lead result in comprehensive results to answer complex problems (Siwi et al., 2022).
Computational linguistics is concerned with a system of words or symbols that can be communicated to a computer. The Linguistics field has dominated in transmitting and interpreting any kinds of symbols into different meanings as it happens in conveying emoji in social media where the social media has become a media to do a social interaction among people via texting to convey any emoji to the particular message (Arafah & Hasyim, 2019; Hasjim et al., 2020). Semiotics, one of the linguistics branches, is also a study of signs and symbols in socio-cultural life in a society where these signs and symbols try to see the relation of humans with groups in their environment (Hasyim et al., 2020). Furthermore, research in this area considers the mathematical and logical characteristics of natural language and develops algorithms and statistical processes for automatic language processing. It is different from the symbols used in the literature field where the symbols can create many different opinions in the reader's mind, while the computational language creates symbols with inexact data and has its meaning (Afiah et al., 2022).

Nowadays, technological advances have made many aspects of life easier to access. The communication systems and educational system meet the new era as online communication and online learning are very possible to reach (Anggrawan et al., 2019). The most visible connection between people and technology that changed the human lifestyle can be seen in many aspects, including cybercrime, climate crisis, deforestation, biodiversity loss, plastic pollution, and many more (Arafah et al., 2021). Along with the rapid development of modernization, the internet user continues to grow and becomes the most effective medium to communicate with people from different places all over the world (Arafah & Hasyim, 2022). In contrast, the negative impact of the technology era can affect even worse to growing children, for example in the digital democracy, the use of social media to put their thoughts or critics that sometimes use bad and inappropriate words (Arafah et al., 2021). Furthermore, one way to face this problem is to introduce the system of coding and programming to children who cannot be separated from the computer and gadgets.

In programming, several errors greatly affect the final result that will be displayed by the code. One of the most common errors in programming is Syntax Error. This type of error is also found in the use of human language. A syntax error in a language is a mistake in using a language that involves organizing words and phrases that do not make sense. Meanwhile, a syntax Error or grammatical error in programming is an error in writing code in a program that makes the format or information unrecognizable by the computer system so that the computer cannot understand the meaning of the code.

In programming, the slightest writing error can result in a syntax error and affect the final result of the program. As for how a teacher will face trying to teach a subject by understanding the course first, this difficulty in syntax error can also be faced by becoming thoroughly familiar with the content and the structure of data at the very first (Sunardi et al., 2018). In this article, the author will show how writing errors in programming can affect the overall results that will be displayed. The author will use the code used to create a Geographic Information System Website for the Spread of Stunting Disease in Banggai Regency with the K-Means method. Stunting is a condition of growth failure in children in the growth of the body and brain, which is caused by malnutrition for a long time and causes the child to have a shorter stature than normal children of their age and have delays in thinking.

This website is designed to find out how stunting distribution in the Banggai district is based on data obtained from the local health office. This website calculates the distribution of stunting in the Banggai Regency using the K-Means Clustering method. The appearance of this website is designed using the PHP language and K-Means calculations can use the R language with R studio software as well as the PHP language.

This article aims to show how the slightest typo can affect the overall appearance of the existing code. This article will show a web view of the correct code as well as code that has a typo in comparison to how a typo can affect the final output.

II. LITERATURE REVIEW

A. Computational Linguistics

Computational linguistics is concerned with understanding language from a computational perspective and constructing artifacts that are useful in processing and generating language (Schubert, 2014). This branch of linguistics is an idea from Noam Chomsky which aims to initiate a particular language to produce a program. An understanding of computational language can also provide insight into thinking and intelligence (Schubert, 2014). Computational linguistics deals with word systems that can be communicated and understood by computers, which can be translated into machine code (Syafar & Febrina, 2019).

Computational linguistics has several theoretical goals that include a grammatical and semantic framework to characterize the language that allows syntactic and semantic analysis to be conducted with computations; discovering processing techniques and learning principles that take advantage of the language's structural and distributional properties; and the development of cognitively plausible and neuroscientific computational models of how language processing and learning can occur in the brain (Schubert, 2014).

Computational linguists are interested in providing computational models of various types of linguistic phenomena, which are "knowledge-based" or "hand-crafted", as well as "data-based" or "statistical" or "empirical". Work in this field is usually motivated from a scientific perspective as one tries to provide computational explanations for linguistics and it can also be one's desire to provide working components of speech or natural language systems (Sproat, 2005; Stern, 2018).
Computational linguistics contains about how processes work with language and linguistics. In computational linguistics, the mechanism for selecting word senses provides an interpretation of what a word with a linguistic structure means, and what that linguistic structure is for. Linguistics in its mainstream form is talking about algorithmic processes, which means about the processes, or in a more comprehensive sense not only of Chomsky's performance but of computing as a generic abstraction (Jones, 2007).

The growth of computational linguistics, as well as natural language information processing, is increasingly being conducted by people with computational rather backgrounds than linguistic backgrounds. In works that study machine learning, more training is needed in mathematics than in linguistics (Jones, 2007; Ibm.com, 2020).

B. Syntax Error in Language

Syntax errors in language use occur when a word and phrase is formed but has a meaning that does not make sense. Syntax shows how a sentence is formed or structured to be a word, which can be misinterpreted or misconstrued. An example of how syntax can affect the context and meaning of phrases or sentences can be seen below (Akorbi, 2021):

- The child cried loudly.
-令 the child loudly cried.
- Lowly, the child cried.

A syntax error, which may seem trivial, can drastically change the meaning of a phrase or sentence as a whole. Some common syntax errors are incomplete sentence structure, subject-verb errors, improper use of conjunctions, incorrect use of prepositions, etc. (Ha'iz et al., 2018). The most common syntax errors that need to be paid attention to are the improper use of commas and the use of sentence fragments (Akorbi.com, 2021). As it is already known that the shift of syntax can cause the changing of meaning, it has to be more careful. The changing meaning of a language is also caused by the shift of lexicons that changed all the cultural values of a language. In the worst case, a language can be extinct if the system turns shifted many times (Takwa et al., 2022).

According to Basri et al. (2013), the syntax was very essential to be understood by the students in learning a language. Another study conducted by Zughoul (2002) showed that errors in noun phrases and verb phrases were frequently conducted by the learners. The most frequent noun phrase errors were in the use of articles, ordinals, and quantifiers (Yuliana, 2017).

C. Syntax Error in Programming

In programming, several errors often occur, one of which is Syntax Error. Syntax error or grammatical error is an error in writing code in a program that makes the format or information unrecognizable by the computer system so that the computer cannot understand the meaning of the code (Sari, 2022). A syntax error can occur when the grammatical rules of the programming language being used are not followed by the person who wrote the program, e.g. if a keyword is misspelled or the author does not put a semi-colon in the appropriate places (Solo, 2020). Syntax error can be a major obstacle for novices and will slow down their progress (Denny et al., 2014). In this kind of error, the compiler finds something wrong with the program. The first step in the debugging process is to fix syntax errors that occur because the program will not run properly if this error is not immediately resolved (Solo, 2020). Syntax errors are one of the main reasons why beginners in this field cannot master programming, due to their inability to apply valid syntax rules when writing programs (Plonka et al., 2015; Mase & Nel, 2022).

Examples of errors that fall into this category are writing commands that do not exist, forgetting to write square brackets, round brackets, and semicolons, misspelling variables, or other errors when writing a programming language (Paskalina, 2021). For compiled languages, a syntax error will result in a compiler message that will usually point to the wrong program line. An example message is “Line 23: Missing semicolon” (Alzahrani & Vahid, 2021).

Research conducted by Denny et al. (2011) explores how the frequency of students who experience compilation errors when writing program code is relatively short. In the study, there were about 70% of students experienced four or more syntax errors in a row even though the compiler output had been shown to them. This syntax error is an error that can be a significant barrier to student success in mastering programming (Denny et al., 2011). An example that shows how serious is this problem involves a student who spent almost 2 hours trying to test whether the sum of two numbers is even or odd (Denny et al., 2014). These kinds of problems are hard to avoid for the students so their willingness to learn from their mistakes is very important. The positive mind of the students to not protract from the problems and arise themselves with self-concept will create positive energy and result in a high motivation to learn better (Arafah et al., 2020). In line with this, students will achieve a learning method themselves that can come both from the student themselves and the environment (Mokoginta & Arafah, 2022). Some of the most frequent syntax errors that beginning coders may write include missing or unmatched parentheses that happen when one end of the parentheses is missing in the code, undeclared or misspelled variables, unmatched or missing quote (’ or “), incomplete or misspelled return statement, and missing semicolon (Woz-u.com, 2021). As a result, the ability is strongly needed by a coder to make a program by putting the language used that has been designed for a coding program, as well as an author who puts his thought through a literary work (Mutmainnah et al., 2022). The similarity between computational linguistics and literature can be seen through the language used because a literary work will be interesting using such a certain beautiful language while computational linguistics can only be done by using such a certain code (Asriyanti et al., 2022).
D. Stunting

Stunting is a chronic nutritional problem that occurs in toddlers, which is caused by a lack of nutritional intake in the long term and due to food intake that is not following nutritional needs. According to UNICEF (United Nations International Children’s Emergency Fund), stunting is the percentage of children aged 0 - 59 months, with a height below -2 (moderate and severe stunting) and -3 (chronic stunting) as measured by the Multicentre Growth Reference Study or the median standard deviation of child growth standards from WHO (World Health Organization) (Indonesia Government, 2020; Anita et al., 2021).

Besides stunted growth, stunting is associated with less than optimal brain development, which can lead to poor mental and learning abilities, as well as poor school performance. Stunting and other conditions are associated with malnutrition and can also be considered a risk factor for diabetes, hypertension, obesity, and death from infection (Lifestyle.kompas.com, 2017; Anita et al., 2021). Poor nutritional status in pregnant women and infants is the main factor causing toddlers to experience stunting. Several factors cause stunting, namely inadequate maternal knowledge, recurrent or chronic infections, poor sanitation, and limited health services.

Some of the symptoms of stunting in children are children who are shorter for their age, body proportions tend to be normal but children look younger/smaller for their age, low weight for their age, and delayed bone growth (Aladokter.com, 2020).

E. Geographic Information System

A geographic Information System (GIS) is a computer-based system that is used to store and manipulate geographic information. A Geographic Information System (GIS) has been designed to collect, store, and also analyze objects and phenomena that present geographic location as an important or critical characteristic to be analyzed.

Geographic Information System (GIS) is formed from three main elements, namely system, information, and geography. In GIS, some elements are most emphasized, namely information and geography. Where the information element describes places, knowledge about the position of a place, and provides information about a position that you want to know. While the geographical element explains that all the information needed is located on the surface of the earth (Prahasta, 2002; Prahasta, 2014).

F. GIS Website for Stunting Distribution in Banggai Regency with the K-Means Method

This website is the first author’s final project to earn a bachelor’s degree in Computer Science. This website was created to show how the distribution of stunting under five in the Banggai Regency with data obtained from the Local Health Office. The data is then processed using the K-Means Clustering method to determine the regional division. The website was built using PHP (Hypertext Preprocessor) to get a dynamic display and connect to the database. The database used to store this web data is MySQL via PHPMyAdmin and uses OpenLayers to display maps on the website. Meanwhile, the K-Means method is calculated automatically on the website using PHP and also uses the R language in the Rstudio application for separate calculations.

PHP (Hypertext Preprocessor) is of programming language in the form of a script that is placed on the server and processed on the server (Prihatna, 2005). PHP is designed to build a dynamic web. This means PHP can form a display according to demand and has good abilities in several things, such as mathematical calculations, in terms of e-mail network information, and regular expressions. PHP is also able to be used as an interface with the database properly and supports various database servers such as MySQL, ORACLE, Sysbase. Meanwhile, OpenLayers is a JavaScript-based client application to displays map data on a web browser and do not depend on the web server used. OpenLayers implements the JavaScript API used to build this website (Laksmita et al., 2012). While the R language is the language used in statistical computing which was first developed by Ross Ihaka and Robert Gentleman at the University of Auckland New Zealand which is an acronym for the first names of the two authors. This language has different rules and syntax from other programming languages which makes it unique compared to other programming languages (Rosidi, 2019).

III. Methodology

The study was conducted with a qualitative method. Qualitative research is research conducted in certain settings that exist in real life (natural) to investigate and understand phenomena: what happened, why it happened, and how it happened (Fadli, 2021). According to Denzin and Lincoln (1994), qualitative research is research using a natural setting to interpret a phenomenon that occurs and is carried out by involving various existing methods (Fadli, 2021). Two types of data that are often used in research are primary and secondary data. Primary data means that the data already exists and was collected by the researcher, while secondary data means that the researcher gathers the data from another source as the supporting data (Purwaningsih et al., 2020).

This study will only use the primary data obtained from PHP code to create a web display of Stunting Distribution in Banggai Regency using the K-Means Method, as well as PHP and R codes to calculate regional division using the K-Means Clustering method. The method of data collection will be done by testing the program code to create a web display and calculating K-means, to see and compare the results that are issued if there is a typo or not. The data used in this article are the results of the first author’s thesis research to get a bachelor’s degree in the Computer Science field.
IV. FINDINGS AND DISCUSSION

On the website that is used as a data source, the K-Means method is used to process the stunting data that has been obtained to be divided into 3 groups and marked with 3 different colors. As for these three groups, namely, groups of areas with low, medium, and high risk.

Meanwhile, the web display that is built is divided into two, namely the user section and the admin section. On the user side, several activities can be carried out by the user, namely viewing the distribution map where the user can choose whether to view by year or by available group, view general information about stunting, view sub-district data, and view information from the sub-district by clicking on click the district area on the map. While on the admin side, the admin can input or edit all the information displayed in the system.

On the initial page of the web display, the user will be shown how the stunting distribution map depicts Banggai Regency. In addition, users can also view other information such as the definition and characteristics of stunting, child body standards, causes of stunting, as well as the impact and prevention of stunting. If the existing code does not have a typo, then the web will display all existing components.

However, if there is a slight typo in the code, it will result in a different appearance on the web. For example, if there is a typo in the background used, then the background cannot appear on the web display. Meanwhile, the code displays the map division with three colors, namely green, yellow, and red, which represents the division of the area with stunting toddlers in the low, medium, and high categories. However, if there is a slight typo in the code, the website will not be able to display the division of the three categories. This can be seen in Figure 1.

In Figure 1, it can be seen that the image on the right can display all components because the code that is executed does not have a typo, so the computer can translate the code correctly. While the two images on the right can't fully display the web part because there is a typo, which causes the computer to not understand the intent of the code. The comparison of the two codes can be seen as follows:

```html
With typo (background):
<body class="animation">
<div class="page-wrapper">
  <div class="background">
    <div class="transbox">
      <center><p><B>PETA PERSEBARAN PENYAKIT STUNTING DI KABUPATEN BANGGAI</B></p></center>
    </div>
  </div>
</div>
<!-- END WELCOME-->  
```

```html
Without typo (background):
<body class="animation">
<div class="page-wrapper">
  <div class="background">
    <div class="transbox">
      <center><p><B>PETA PERSEBARAN PENYAKIT STUNTING DI KABUPATEN BANGGAI</B></p></center>
    </div>
  </div>
</div>
<!-- END WELCOME-->  
```
From the code snippet, it can be seen that there is a typo on line 3 for the background code, which causes the background display to not appear. The `<div class="background">` code only has a typo in the word "class" but can make the computer unable to read what this code means and result in no background appearing on the web display. Meanwhile, in the code to display the map category division, there is a typo in the code `$query = mysqli_query($conn, "selrvt * from tb_region");` which makes the computer unable to process the code.

Then on another page for users, which is the child's body standard page, it will show a standard table of the child's body from 24 to 60 months. The table will show the child's body standards from -3SD, -2SD, -1SD, Median, +3SD, +2SD, and +1SD. On this page, the user can choose to view standard body information for boys and girls and can view by age and height. However, if there is a little typo, it will make the web unable to display the entire table.

![Figure 2 Display of Map Category Division With Error Typing in Code](image)

In Figure 2, it can be seen that the image on the right shows the standard table for children's bodies, where the user can choose to view the standard data for boys and girls based on age and height. While the two images on the right cannot display the entire table because there is a typo, which causes the computer to not understand the intent of the code. The comparison of the two codes can be seen as follows:

```
With typo (standard body):
<?php
    if(strtolower($judul) == 'standar badan anak') {
        include 'standar_badan_anak.php';
    } else {
        echo $data['isi'];
    }
}
```

```
Without typo (standard body):
<?php
    if(strtolower($judul) == 'Standar Badan Anak') {
        include 'standar_badan_anak.php';
    } else {
        echo $data['isi'];
    }
```

On the admin page, there is a page to recalculate the division of categories using K-Means clustering. On this page, K-Means are calculated automatically using the PHP language. If there is no typo, the web will display a page to update the existing cluster with the K-Means Clustering method. However, if there is a typo even if it is only one letter, then the page will be blank and not display anything.
In Figure 3, it is clear how the page changes when there is a typo in the code used. When there is no typo, the page can load a table containing the results of the cluster calculations using K-Means Clustering. However, when a typo occurs, the page will be blank so that it displays anything. The comparison of the two codes can be seen as follows:

**With typo (K-Means):**
```
$query = mysqli_query($conn, "select * from tb_tahun");
    while($stahun = mysqli_fetch_array($query)){
        $data =
        mysqli_num_rows(mysqli_query($conn, "select id_tahun from tb_data_kecamatan where id_tahun=".$stahun["id_tahun"]")));
        echo "
```

**Without typo (K-Means):**
```
$query = mysqli_query($conn, "select * from tb_tahun");
    while($stahun = mysqli_fetch_array($query)){
        $data =
        mysqli_num_rows(mysqli_query($conn, "select id_tahun from tb_data_kecamatan where id_tahun=".$stahun["id_tahun"]")));
        echo "
```

From the code, it can be seen that there is only a minor typo in the code `$query = mysqli_query($conn, "select * from tb_tahun");` on the word "query" to "queryy". However, this is fatal to the web display that should display the clustering results from the K-Means calculation.

Meanwhile, for the calculation of K-Means Clustering using the R language through the Rstudio software, the correct program will produce a table that has divided the data. However, if there is an error in typing, the program will immediately stop processing.

From all the sample data that has been displayed, it can be seen that the typo greatly affects the running of the code. When a code experiences a typo or syntax error, it will cause the computer to not understand the meaning of the code. Thus, the computer will not be able to display what the codes want.

**V. Conclusion**

A syntax error is an error that often occurs in the use of linguistics, both in language and linguistics branches. One of the linguistic branches that can be affected by syntax error is computational linguistics, especially in programming code. In programming code, even the smallest typo can affect how the result will display in the code. When there is a typo in the program, then the computer will not be able to understand the code, and made the computer cannot display it.
From the explanation above, it can be concluded that the way the code is written is very influential on the results that will be displayed by the code. Code in dominant programming uses simple words in English. However, this sometimes becomes an obstacle for beginners who tend to make mistakes when typing the code. If someone makes a writing error in the code, this will cause a syntax error and will greatly affect the results obtained.

When making an error like this, the computer will not be able to understand the intent of the code which results in the computer not being able to display the results that the author or coder intended. Errors like this are easier to fix than other errors in programming, it’s just that precision is needed to find the part of the name that contains typos, especially in long programs.

REFERENCES


Burhanuddin Arafah obtained his Ph.D. degree in English Literature at the University of Newcastle, NSW Australia in 2003. He earned his Master's degree in American Literature at Gadjah Mada University (UGM) Yogyakarta of Indonesia in 1995, and his Bachelor's degree in English Literature at Hasanuddin University (UNHAS) in 1988.

He is currently a full Professor in English Literature at the English Literature Study Program, Faculty of Cultural Sciences of Hasanuddin University (UNHAS) Indonesia. He has published four books in English language and literature and more than 50 research articles ranked in international journals published in the English language. His areas of interest are English literature, language education, and cultural studies.

He was the Dean of the Faculty of Cultural Sciences of Hasanuddin University from 2009 to 2017, and
currently, he is actively involved at the National Accreditation Board-Higher Education, Ministry of Education and Culture of the Republic of Indonesia for his position as Assessor.

Syanuwalini Syafruddin obtained her Bachelor's degree from The State University of Makassar (UNM), majoring in English Language Teaching at the Faculty of Language and Art, UNM in 2007. She is currently an English teacher of Madrasah Tsanawiyah (MTs) in Makassar while pursuing her study in English Language Studies Program, majoring in Linguistics at the Postgraduate Program of the Faculty of Cultural Sciences, Hasanuddin University.

Jumardin Muchtar has got S.S. (Bachelor of Letters) at Makassar Islamic University, Makassar City in 2020. His scholarly interest includes English Literature, English Language Education and Linguistics. He has published several articles, such as Grammatical Errors in English Informal Essays by The Sixth Semester of English Literature Students Vol. 4 No. 1 (2021): March, Local Folklore and Reading Comprehension of Eighth Grade Students at SMP N 2 Bau-Bau, An Analysis of English-Indonesian Translation Shift in Little Match Girl Short Story Vol. 1 No. 1 (2021) and The Analysis of Stab Subtype with Comparative Study between English and Buginese Construction Vol. 2 No. 1 (2022). He is currently undergoing a master's degree in English Language Studies Program, majoring in Linguistics at the Postgraduate Program of the Faculty of Cultural Sciences, Hasanuddin University in 2021-present.

Puji Ayu Lestari obtained her Bachelor's degree from Hasanuddin University majoring in the English Department at the Faculty of Cultural Sciences in 2013. Currently, she is a student in English Language Studies Program, majoring in Linguistics at the Postgraduate Program of the Faculty of Cultural Sciences, Hasanuddin University.