Chapter 25

Descriptive Data Analytics on Dinesafe Data for Food Assessment and Evaluation Using R Programming Language:
A Case Study on Toronto's Dinesafe Inspection and Disclosure System

Ajinkya Kunjir
https://orcid.org/0000-0001-7634-4115
Lakehead University, Canada

Jugal Shah
Lakehead University, Canada

Vikas Trikha
Lakehead University, Canada

ABSTRACT

In the digital era of the 21st century, data analytics (DA) can be highlighted as 'finding conclusions based on observations' or unique knowledge discovery from data (KDD) in form of patterns and visualizations for ease of understanding. The city of Toronto consists of thousands of food chains, restaurants, bars based all over the streets of the city. Dinesafe is an agency-based inspection system monitored by the provincial and municipal regulations and ran by the Ministry of Health, Ontario. This chapter proposes an efficient descriptive data analytics on the Dinesafe data provided by the Health Ministry of Toronto, Ontario using an open-source data programming framework like R. The data is publicly available for all the researchers and motivates the practitioners for conveying the results to the ministry for betterment of the people of Toronto. The chapter will also shed light on the methodology, visualization, types and share the results from the work executed on R.

INTRODUCTION

Data science can be generally explained as a lifecycle of data gathering, preparation, transformation and pattern generation to achieve milestones. The data collected has many dependencies starting from time, space and complexity. Big Data and Advanced deep learning are the two rapidly growing areas of research and advances. Hence, there is a need to construct or modernize the process that resolves and addresses current challenges in the data development and deployment cycle. Several concepts about data science have been derived by the multinational companies dealing with massive chunks of data in everyday life. Data measurable in TeraBytes (Tb's), Petabytes (Pb's), and Zettabytes are being generated by social media sites such as Facebook, Twitter, and LinkedIn daily. The data collected is unstructured or semi-structured as it comprises images, audio, documents and all other media. The goal is a business outcome rather than improving a measure of accuracy on a specific analytic model. Operational Big Data analytics or systems implementing Operational Big Data Analytics (BDA's) introduce a plethora of difficulties and challenges due to the data distribution, novel approach(s) and partnership with other multiple organizations that are using cloud services (Nancy W. Grady et.al, 2017). Where Big Data is booming in the field of analytics, the decision-making flows along with it. There are four types of analytics to start with: Descriptive Analysis, Diagnostic analytics, Predictive Analytics, and Prescriptive analytics. The more complicated the analytics is, the more value it holds. For instance, to deal with our data in this research, the authors will be going forward with descriptive analytics. Descriptive analytics deal with the question 'What Happened? Descriptive analytics mostly juggles data from multiple sources to deduce insights into the past. The patterns generated simply signal that is something right or wrong, without a proper explanation. With the 'Dinesafe' data, the authors are going to find out 'what happened with what?' for the current and past few years (2017 & 2018). The following questions are to be investigated in this research

- What was the severity of food establishment in 2017, 2018 and 2019?
- How has the progress been from 2017 to 2019?
- How many Food establishments types have a minimum number of inspections = '1/2/3'?
- The ratio of a pass, conditional pass and Closed for all food establishments in 2017, 2018 & 2019?

There may be many more questions arising by looking at the dataset, but we are just going to showcase the ones that have a high impact on the viewers.

Existing System

The current work executed on Dinesafe is reported and embedded in the Toronto Health safe website as an application with inputs. The work is illustrated in the form of a map representation pointing at all the current food establishments according to their result flags such as Pass(In Green), conditional pass (Yellow) and closed (In red). The web application can be queried by the end users and return outputs based on the inputs they feed to the fields. Currently, the app can accept 'Postal code' and 'Food establishment Name' from the end-users and represent the output on the map itself. The existing system is worthy of producing the immediate results to the users but lacks prediction, descriptions and comparative analysis between the years. The lacuna of the current system can be filled in by adding informative layers of technology to the Dinesafe data, elaborated in the next few sections.

21 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/descriptive-data-analytics-on-dinesafe-data-for-food-assessment-and-evaluation-using-r-programming-language/264325

Related Content

Proximate Breast Cancer Factors Using Data Mining Classification Techniques
Alice Constance Mensahand Isaac Ofori Asare (2019). *International Journal of Big Data and Analytics in Healthcare (pp. 47-56).*
www.irma-international.org/article/proximate-breast-cancer-factors-using-data-mining-classification-techniques/232335

Contour Reconstruction: 2D Object Modeling
(2017). *Probabilistic Nodes Combination (PNC) for Object Modeling and Contour Reconstruction (pp. 132-177).*
www.irma-international.org/chapter/contour-reconstruction/180357

Virtual Kernel Discriminative Dictionary Learning With Weighted KNN for Video Analysis
Ben-Bright Benuwa (2022). *International Journal of Data Analytics (pp. 1-19).*
www.irma-international.org/article/virtual-kernel-discriminative-dictionary-learning-with-weighted-knn-for-video-analysis/297521

Multiobjective Transportation Problem Using Fuzzy Decision Variable Through Multi-Choice Programming
Gurupada Maityand Sankar Kumar Roy (2018). *Intelligent Transportation and Planning: Breakthroughs in Research and Practice (pp. 866-882).*
www.irma-international.org/chapter/multiobjective-transportation-problem-using-fuzzy-decision-variable-through-multi-choice-programming/197165

Relationships Between Out-of-School-Time Lessons and Academic Performance Among Adolescents in Four High-Performing Education Systems
David R. Litz, Shaljan Areepattamanniland Scott H. Parkman (2022). *Applying Data Science and Learning Analytics Throughout a Learner's Lifespan (pp. 27-55).*
www.irma-international.org/chapter/relationships-between-out-of-school-time-lessons-and-academic-performance-among-adolescents-in-four-high-performing-education-systems/301853