Bharath
INSTITUTE OF HIGHER EDUCATION AND RESEARCH
(Declared as Deemed-to-be University under section 3 of UGC Act, 1956)
(Vide Notification No. F.9-5/2000 - U.3, Ministry of Human Resource Development, Govt. of India, dated 4th July 2002)
Accredited by NAAC NATIONAL ASSESSMENT AND ACCREDITATION COUNCIL
Phone : 044-22290742 / 22290125 . Telefax : 044-22293886
Website : www.bharathuniv.ac.in
173, Agaram Road, Selaiyur, Tambaram,
Chennai - 600 073. Tamil Nadu.

Ref No.SMS-2018-O-26 Date: 07/11/2019

TO

Mrs. K. Hemalakshmi,

Associate Professor/ CSE,

BIHER.

BHARATH INSTITUTE OF HIGHER EDUCATION & RESEARCH (Declared as Deemed to be University U/S 3 of UGC Act 1956)
07 NOV 2019
OFFICE OF THE DEAN-RESEARCH
Selaiyur, Chennai-600 073. INDIA

Thro: Concern Head of the Department

Greetings!!!

We are happy to announce that the Research Advisory Committee has approved your proposal for Seed Money Scheme-2018 which was presented by you. You are requested to complete the proposal and send the progress report to the Dean Research in the prescribed time period.

Title of the Project: A new framework for statistical thinking in times of Big Data and Digital Economy.

Seed Money Amount: Rs.1, 00,000/- (Rupees One Lakh Only)

Approved on: 04/11/2019

Payment details:

Voucher No.26 **Dated: 08/11/2019**

With Regards

Dean-Research

Bharath University

SELAIYUR, CHENNAI - 600 073, TAMIL NADU, INDIA.

CASH / PAYMENT VOUCHER

Date 08/11/19

V.No: 026

Debit _____ Amount

Rs. 1,00,000/-

PAID TO Mrs. K. Hemalakshmi

RUPEES one lakh only.

TOWARDS Seed Money Scheme - 2018.

Authorised by **Finance Manager** **Cashier/Accountant** **Payee's Signature**

PROPOSAL SUBMISSION

1. **Details of Principal Investigator**

Name	: K.Hemalakshmi
Designation	: Assistant Professor
Highest Qualifications	: M.C.A.,M.Phil.,(Ph.D)
Department	: Computer Science
E-mail	: hemalakshmi.mca@bharathuniv.ac.in
Contact no	: 9790156157
Date of Joining	: 13.07.2015

2. **Details of Co - Principal Investigator**

Name	: Dr.A.Muthukumaravel
Designation	: Dean Faculty of Arts and Science, Prof and Head Department of MCA.
Highest Qualifications	: Ph.D.
Department	: Computer Science
E-mail	: muthukumaravel.mca@bharathuniv.ac.in
Contact no	: 9500137273
Date of Joining	: 12.07.2013

Technical details

1. Introduction

The interest in big data is growing exponentially in today's society. Commercial insights, government initiatives and even research calls, all seem to be focused on exploiting the potential of technology to capture and analyze massive amounts of data in increasingly powerful ways. Big data, that is, data that are too big for standard database software to process, is everywhere. For some, big data represents a paradigm shift in the ways that we understand and study our world, and at the very least it is seen as a way to better utilize and creatively analyze information for public and private benefit.

The concept of big data "refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" (Manyika et al., 2011, p.1). Additionally, big data is often associated with key characteristics that go beyond the question of size, namely the 5 Vs: volume, velocity, variety, veracity and value (Storey & Song, 2017). Big data is dispersed among various platforms that operate with different standards, providers and degrees of access (Ferguson, 2012). For example, a lot of work in big data focuses on Twitter, the blogosphere, and search engine queries. All of these activities are not undertaken equally by the whole population, which raises concerning issues around the question of whose data traces will be analyzed using big data.

There are also a number of practical issues related to working with big data. These include, among others, issues we cannot afford to ignore, such as implications for the training of future teachers regarding handling and analysis of big data.

Due to the fact that big data has recently become mainstream in many research fields, including education, it is important to discuss and answer the following questions:

1. In order to function effectively in a society driven by big data and digital economy, what are the necessary processes of statistical thinking required to handle big data?
2. How can we revise current curriculum frameworks of statistical thinking to incorporate big data for the digital economy?
3. How can we incorporate core ideas of big data for the digital economy into the high school curriculum?
4. What are plausible instructional activities (exemplar applications) for teaching the fundamental ideas of statistics while fostering statistical thinking for big data and the digital economy?

2. Review of status of Research and Development in the subject

In this digital era, statistical thinking processes do not follow the Problem-Plan-Data-Analysis-Conclusion (PPDAC) cycle (Wild & Pfannkuch, 1999) anymore, due to the shift in

the way we work with data set by the arrival of big data analytics. In fact, the PPDAC cycle is a question-then-answer research method, focused on data gathered for a purpose using planned processes, chosen on statistical grounds to justify certain types of inferences and conclusions. However, in times of big data, this is actually a weakness of the PPDAC cycle, because most of the data available is opportunistic (happenstance or "found") data (including "big data"): huge amounts of data already collected by others and hosted somewhere. Nowadays, many companies have data teams exploring large sets of raw opportunistic data, looking for new connections and identifying significant correlations, while refining their analysis until they arrive at valuable understandings. This approach reverses the question-then-answer process of the PPDAC cycle. It starts with strong, data-first answers, and then works backward to find the questions that should have been asked.

must acknowledge that any up-to-date framework for statistical thinking must be designed giving consideration to these criticisms to the PPDAC cycle. By doing so, we came up with the following framework to describe how a person engages in statistical thinking while handling big data. The proposed framework understands statistical thinking as a cognitive process comprised of the following five phases:

Patterns and relationships from data: Look for patterns and relationships within the data, based on a particular interest.

Questions: Pose critical and worry questions, in order to find plausible explanations to the patterns and relationships found.

Objectives: Set objectives related to the posed questions, in order to analyze the data.

Data mining: Re-examine the data in the light of the objectives, explore the old and new data sources, or introduce new variables for consideration. Data mining can be data-oriented, explanation-oriented, or future- oriented.

Understanding and/or designing: Provide ideas for new activities, based on the understanding of the past, and design plans and strategies for the future, based on the results from the data mining.

A detailed explanation of this framework for statistical thinking (González, Isoda & Araya, 2019) was submitted for publication in the Statistical Education Research Journal (SERJ).

Examplar Application of The Framework: Aging Population Issues In Apec Countries

For the purpose of exemplifying this framework, let us suppose that we are interested in exploring issues related to population ageing, which is a concerning issue in many societies, such as the Japanese. So, we may check the worldwide trend of web searches for terms such as "social security" and "nursing home", focusing on APEC countries.

For this example, we will use the website "Google Trends" (https://trends.google.com), which is an open- access online platform for big data that enables creative discovery from information.

2.1 International Status: NIL

2.2 National Status: NIL

3. Progress/achievement so far,

a) Reference papers was collected.

b) Literature survey was studied.

c) Proposal work has been started in the wind turbine based PMSG fed three phase inverter for grid connected system.

4. Work Plan:

4.1 Methodology:

The main objectives are as follows, Patterns and relationships from data

Using "Google Trends", we looked for patterns and relationships within the data hosted in the platform, based on our interest in aging population issues on APEC countries. Then, by checking the worldwide trend of web searches for terms related to aging population, such as "social security" and "nursing home", we will be able to identify possible patterns and relationships in the data regarding these terms. Figure 1 shows the evolution of the worldwide search trends for "social security" and "nursing home" in the past 15 years.

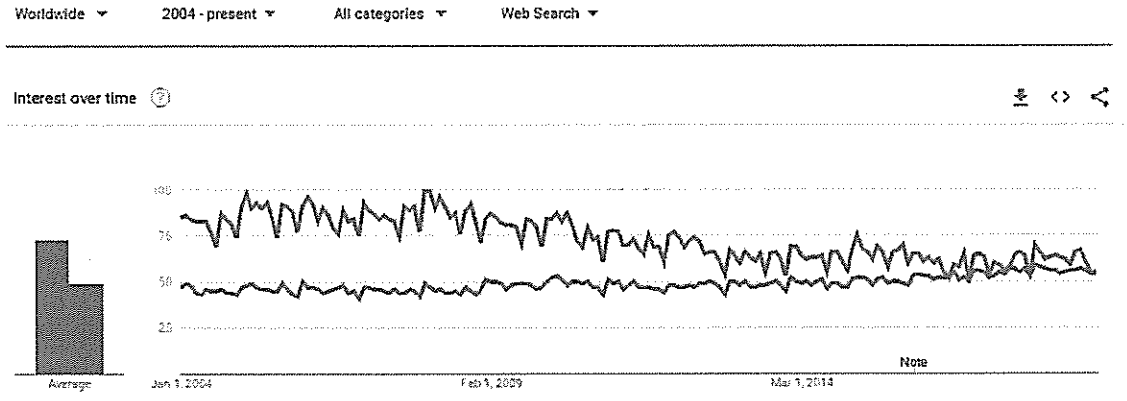


Figure 1: Evolution of the worldwide Google search trends for "social security" (in blue) and "nursing home" (in red) in the past 15 years.

Now, let us focus on the online search trends for the terms "social security" and "nursing home", focusing on APEC countries. In the top row of Figure 2, we can see that Japan,

Canada and the US were APEC countries in which, in the last year, online searches for the term "nursing home" were higher in comparison to the term "social security". On the other hand, in the bottom row of Figure 2, we can see that Chile, Peru and the South Korea were APEC countries in which, in the last year, online searches for the term "social security" were higher in comparison to the term "nursing home".

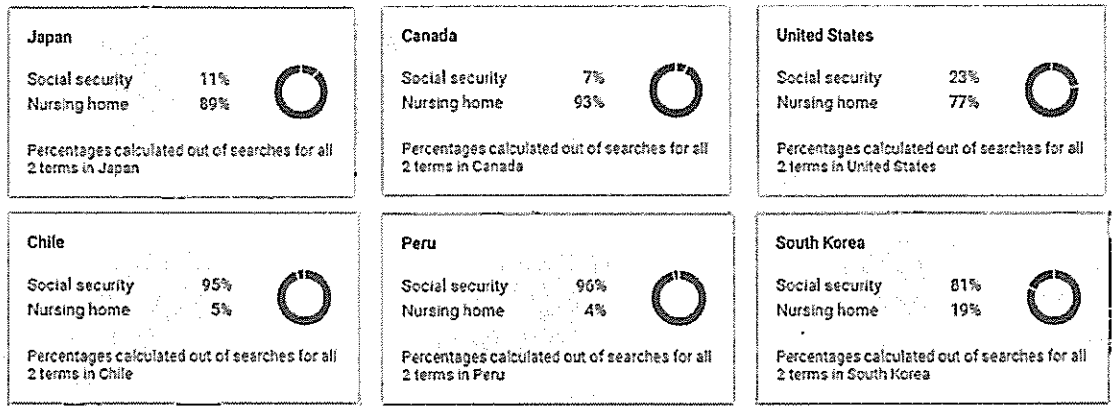


Figure 2: Percentage comparison of Google searches for the terms "social security" (in blue) and "nursing home" (in red) in six APEC countries in 2018.

Now, from the posed questions, we are able to set clear objectives to address. In fact, each objective should be associated to at least one question. In this example, some objectives stemming from the questions above are the following:

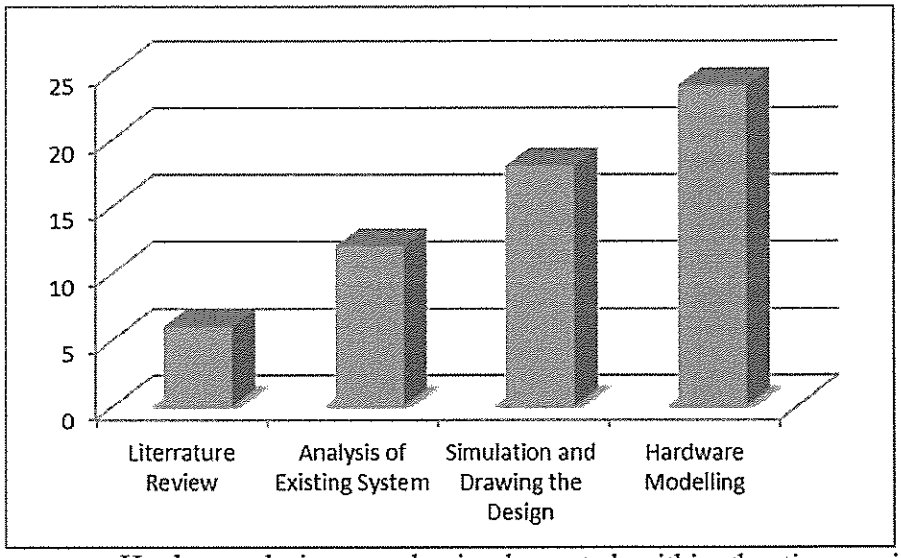
1. To look for and identify the reasons why some APEC countries seem to show considerably more or less interest on "nursing home" than on "social security"
2. To predict the trends of web searches for the terms "nursing home" and "social security" in APEC countries in the next decade.
3. To determine the behavior of related queries (e.g., "nurse" for countries showing more interest on "nursing home", and "tax" or "pension" in countries

4.2 Time Schedule of activities giving milestones through BAR diagram.

Work plan (including detailed methodology and time schedule)

Sl. No.	Activity / Milestone	1st Year		2nd Year	
1.	Literature review	**1-6**			
2.	Analysis of existing system		**7-12**		
3.	Simulation and Drawing the design			**13-18**	
4.	Hardware modelling				**19-24**

4.3 Expected outcome within the time period of Seed Money Scheme



a) Prototype Hardware design can be implemented within the time period of Seed Money Scheme.

b) For a real time HRES field work can be done within the time period of Seed Money Scheme.

5 Suggested Plan of action stating the name of funding agency where the project will be communicated for financial support within the time period of project.

Nil

6 **Bibliography:** Nil

7 List of Projects submitted/implemented by the Investigators (Separate for Pi and Co-PI)

Nil

7.1 Details of Projects submitted to various funding agencies:

Sl. No.	Title	Cost in lakhs	Month of submission	Role as PI/ Co-PI	Agency	Status
	NA	NA	NA	NA	NA	NA

7.2 Details of Projects under implementation

Sl. No.	Title	Cost in lakhs	Duration	Role as PI/ Co-PI	Agency
	NA NA	NA	NA	NA	NA

7.3 Details of Projects completed during the last 5 years

Sl. No.	Title	Cost in lakhs	Duration	Role as PI/ Co-PI	Agency
	NA	NA NA	NA	NA	NA

8 List of publications published by the Investigators, if any:

a) Co - Principal Investigator

S.No	Author names	Title of paper	Name of Journal	Vol (issue)	Page no.	Year
1.	Dr.A.Muthukumaravel, Purushothaman S	CONCURRENCY CONTROL IN CAD USING FUNCTIONAL BACK PROPAGATION NEURAL NETWORK	International Journal of Computer Science and Information Security (IJCSIS)	(Vol.9 No.2) 2011	168-174	2011
2.	Dr.A.Muthukumaravel , Purushothaman S	IMPLEMENTATION OF LOCALLY WEIGHTED PROJECTION REGRESSION NETWORK FOR CONCURRENCY CONTROL IN COMPUTER AIDED DESIGN	International Journal of Advanced Computer Science and Applications (IJACSA)	(Vol.2 No.8)2011	46-50	2011
3.	Dr.A.Muthukumaravel, Purushothaman S	DEVELOPMENT OF FUZZY LOGIC FOR CONCURRENCY CONTROL IN COMPUTER AIDED DESIGN DATABASE	CIIT International Journal of Fuzzy Systems	(Vol.4 No.8) 2012	310-313	2012
4.	Dr.Muthukumaravel A, Prasanna S., Deepa S	SUPPORTING VARIOUS TECHNIQUES TO OPTIMIZE AND SECURE APPLICATION PERFORMANCE IN A CLOUD COMPUTING SECURITY IN A EFFECTIVE	International Journal of Emerging Technology and Advanced Engineering	Vol.3, Issue. 4, April 2013	778 – 781	2013

		MANNER				
5.	Dr.Muthukumaravel A., Shanthi C., Hema R	ENHANCING INTEROPERABILITY AMONG MOBILE APPS USING LCIM MODEL	International Journal of Emerging Technology and Advanced Engineering	Vol. 3, Issue 5, May 2014	491-497	2014
6.	Dr.Muthukumaravel A., Subhapriya G	OPTIMAL SOFTWARE TESTING STRATEGIES - EXTRACT	CIIT International Journal of Software Engineering and Technology	Vol.5, No.4, April 2015	126-130	2015
7.	Dr.Muthukumaravel A., Mayilvahanan P., Dhasna Rajasekaran	SOFTWARE TESTING - AN INSIGHT INTO PROCESS IMPROVEMENT	CIIT International Journal of Software Engineering and Technology	Vol. 5, No.5, May 2016	166-170	2016
8.	Dr.Muthukumaravel A., Mayilvahanan P., Saranya K	A Robust Digital Image Watermarking is using Wavelet Filters	International Journal of Latest Trends in Engineering & Technology	Vol.2, Issue 3, May 2017	150 - 154	2017
9.	Dr.Muthukumaravel A., Karthick K., P.Jennifer	SSL Backend Forwarding Scheme in Cluster-Based Web Servers	Middle-East Journal of Scientific Research	Volume 20 Number (6), 2018	752 – 755	2018

b) Principal Investigator

S.No	Author names	Title of paper	Name of Journal	Vol (issue)	Page no.	Year
1.	K.Hemalakshmi, S.Surya	A COMPARATIVE STUDY ON CLUSTERING ALGORITHMS FOR CLUSTERING TEXT DOCUMENTS	Business Analytics and Intelligence(ICBAI-2016)	s.no.2.96, page no-85	168-174	2016

9. **Budget**

Sl. No.	Equipment	Quantity	Amount in INR
1	Battery – 12V, 300AH	2	50,000
	Sine Wave Inverter – 1000Watts	1	15,000
2	Consumables (Like, testing tools Charge controller, etc.)	As per requirement	20,000
3	Travel support for the purpose of research work.	---	5,000
4	Contingency	---	5000
5	Others	---	5000
	Total		1,00,000

10. **Name of at least two subject experts from the Institute and one from the outside Institute with their contact details:**
    a) Dr.S.Sivakumar– Principal, Prof and Head, Dept of Computer science, CPA college Bodinayakkanur
    b) Dr.R.Murugaiah – Chairman, S.Veerasamy college of Engineering and Technology ,Pulinagudi

CERTIFICATE FROM THE INVESTIGATOR

Project Title: A New Framework for Statistical Thinking in Times of Big Data and Digital Economy

It is certified that

1. I do hereby agree to submit a complete proposal for financial support to the external funding agency within the time period of SMS-2018

2. I undertake that spare time on equipment procured in the project will be made available to other users.

3. I agree to submit a certificate from Institutional Biosafety Committee, if the project involves the utilization of genetically engineered organisms. I also declare that while conducting experiments, the Biosafety Guidelines of Department of Biotechnology, Department of Health Research, GOI would be followed in to.

4. I agree to submit ethical clearance certificate from the concerned ethical committee, if the project involves field trails/experiments/exchange of specimens, human & animal materials etc.

5. I agree to abide by the terms and conditions of SMS-2018, BIHER, and Chennai.

K. Hemalatha

Name and signature of
Principal Investigator

Name and signature of
Co-Principal Investigator

Dean - Faculty of Arts & Science
Bharath Institute of Higher Education & Research
(Declared as Deemed to be University U/S 3 of UGC Act, 1956)
Chennai-600 073. INDIA

Date: 03.10.2019
Place: Chennai - 73

Forwarded by Head of the Department

Dr.D.KERANA HANIREX
Associate Professor & Head
Dept.of Computer Applications
Faculty of Arts and Science
BIHER, Chennai-600 126

Signature of the Head

PROJECT EVALUATION FORMAT

Recommendation Sheet

Name of the Principal Investigator	Mrs.K.Hemalakshmi
Name of the Co-Investigator	Dr.A.Muthukumaravel
Name of the Department	Computer science
Title of project	A New Framework for Statistical Thinking in Times of Big Data and Digital Economy
Recommendation of the evaluation committee	Yes.
Financial allocation recommended	Rs. 1,00,000 —

Sl. No.	Equipment	Quantity	Amount in INR
1	Battery – 12V, 300AH	2	50,000
	Sine Wave Inverter – 1000Watts	1	15,000
2	Consumables (Like, testing tools Charge controller, etc.)	As per requirement	20,000
3	Travel support for the purpose of research work.	---	5,000
4	Contingency	---	5000
5	Others	---	5000
	Total		1,00,000

Name and Signature of the Research Advisory Committee members with date

Recommended

Dr. P. Naveenchandran

Internal Quality Cell * BIHER * Chennai-73