

# QUESTIONNAIRE

## SURVEY OF OFFICIAL STATISTICS PRODUCTION PAKISTAN (SOS-Pak)

### INSTRUCTIONS:

1. This questionnaire consists of three Parts, PART-1 & 3 is requested to be filled in by any well aware and experienced officer of this organization, whereas PART-2 preferably be filled in by any IT professional of this organization.
2. Options with numeric codes are single answered whereas options with alphabets may attain multiple answers.

### PART - 1

INFORMATION PANEL						IP
IP 1	Organization Name					
IP 2	Respondent Designation			IP 3	BS or Equivalent	
IP 4	Office Phone Number					
IP 5	Organization Strength (Employees of Attach dept. not inclusive)					<input type="radio"/> Less than 50.....1 <input type="radio"/> 250-1000.....3 <input type="radio"/> 50-250.....2 <input type="radio"/> Above 1000.....4
OFFICIAL STATISTICS PRODUCTION INFORMATION						OS
OS 1	Organizations usually record data about different official functions/activities for public/administrative use. (E.g. students' academic scores, Patients details, food stock records, Citizens biodata, etc.)					
	Does this organization collect/record any kind of Data (For Official or Public Use) regularly?				Yes.....1      » Continue No.....2      » Go to <b>OS 9</b>	
OS 2	The data is collected for which purpose?	<input type="radio"/> Statistical Use.....1 <input type="radio"/> Both 1 & 2.....3 <input type="radio"/> Administrative Use.....2 <input type="radio"/> Others Specify.....4				
OS 3	Does this organization produce any kind of Statistical Summary/Report/Publication/Gazette on this data regularly?				Yes.....1      » Continue No.....2      » Go to <b>OS 6</b>	
OS 4	Please list down the title of Summary/Publications/Reports/Gazette, this organization produces.					
	TITLE OF PUBLICATION/REPORT/SUMMARY/GAZETTE etc.				Daily	Weekly
	1)				D	W
	2)				D	W
	3)				D	W
	4)				D	W
	5) Please use a separate sheet for more records or give the website reference here if any.				D	W
OS 5	Does the organization disseminate these products for official/public use?				Yes.....1 No.....2	
OS 6	Does this organization supply data to Statistical Organizations (Provincial/Federal) periodically?				Yes.....1 No.....2	
OS 7	Does this organization have a framework to deal with privacy issues related to data sources?				Yes.....1 No.....2	
OS 8	How much privacy and confidentiality relates to your data?				<input type="radio"/> Low <input type="radio"/> Medium <input type="radio"/> High	
OS 9	Does this organization need to conduct survey to fulfill its data needs?				Yes.....1 No.....2	
OS 10	Does this organization acquire data from Statistical Organizations periodically for its own needs?				Yes.....1 No.....2	
OS 11	There are many data items or data sets, are usually collected by the organizations but do not report in any format, and hence left unreported for planning, policy making, researchers and academia.					
	Does this organization contain unreported data sources?				Yes.....1      » Continue No.....2      » Go to <b>Next Part</b>	

OS 12	Please list down all the data items / data sets being unreported.								
	1)								
	2)								
	3)								
	4)								
5) Please use a separate sheet for more records.									
OS 13	Do you think, from these unreported datasets, some valuable Statistics can be produced for better planning/decision making?				Yes.....1 No.....2				
OS 14	Please rate the reasons of data sources being unreported.				Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	A	These data sets are confidential/Important			1	2	3	4	5
	B	These data sets seem less important/useless			1	2	3	4	5
	C	Data sets are not computerized.			1	2	3	4	5
	D	Data sets are not being stored in a proper database			1	2	3	4	5
	E	There is no demand for such data sets in market			1	2	3	4	5
	F	Organization do not bother to summarize these sources			1	2	3	4	5
	G	Others (Please Specify)			1	2	3	4	5

## PART - 2

Note: This module should preferably be filled in by any IT professional of this organization

### BIG DATA USE IN OFFICIAL STATISTICS

BD

BD 1	How the data is being recorded in this organization?	<input type="radio"/> Electronic.....1 <input type="radio"/> Manual.....2 <input type="radio"/> Both 1&2 .....3		
BD 2	Where the data is being stored? (More than one options possible) <input type="checkbox"/> Webserver.....A <input type="checkbox"/> Local System/PC.....C <input type="checkbox"/> Data Base Management System Software.....B <input type="checkbox"/> Cloud Storage.....D			
BD 3	Classify the following characteristics of the data produced/recorded/stored in this organization.			
	VELOCITY: What is the maximum data inflow time?	<input type="radio"/> Daily <input type="radio"/> Monthly <input type="radio"/> Yearly		
	VOLUME: How much data is being produced in this Velocity?	<input type="radio"/> < 1GB <input type="radio"/> 1-10GB <input type="radio"/> >10GB		
	VARIETY: What kind of data is being recorded?	<input type="checkbox"/> Numeric <input type="checkbox"/> Text <input type="checkbox"/> Graphic <input type="checkbox"/> Others (Specify)_____		
BD 4	<i>Big data is the digital information produced by wide-range of high throughput technologies and tools about human activities, e.g. ATM transaction, phone calls, satellite imagery, weather/traffic sensors etc. Big Data are data sources with a high volume, velocity and variety of data, which require new tools and methods to capture, curate, manage, and process them in an efficient way.</i>  Do you ever hear about Big Data before now? Yes.....1 No.....2			
BD 5	Do you think Big data may support your organizational planning and decision making?	Yes.....1 No.....2		
BD 6	Big Data is the valuable source with respect to,	Not at all	To Some Extent	To Great Extent
	a). Cost saving in production of required statistics	1	2	3
	b). Improves the quality in the production of statistics	1	2	3
	c). Saves time in the production of statistics	1	2	3
BD 7	Does your organization work with Big Data?	Yes.....1    » Continue No.....2    » Go to <b>BD 9</b>		
BD 8	Which Big Data source is being used by this organization.	A. _____ B. _____		
BD 9	Have your organization plan to work with Big Data in future?	Yes.....1 No.....2		

BD 10	Which type of Big Data has potential for your organization?										
	<input type="checkbox"/> <u>Administrative Data (A)</u> E.g. Medical, Educational, Agricultural, demographic, business records etc.	<input type="checkbox"/> <u>Communication/Tracking devices data (C)</u> E.g. Telecommunication data, Call Detail Records, GPS points etc.									
	<input type="checkbox"/> <u>Commercial or transactional records (T)</u> E.g. Credit or debit card transactions, e-payments, Web scraping etc.	<input type="checkbox"/> <u>Behavioral data (B)</u> E.g. Internet Searches, web page views, web clicks online streaming etc.									
	<input type="checkbox"/> <u>Sensors data (S)</u> E.g. Satellite imagery, traffic, weather sensors etc.	<input type="checkbox"/> <u>Opinion records (O)</u> E.g. Social media updates like Facebook, twitter etc.									
BD 11	Does this organization is well IT equipped and have sufficient resources to fulfill its Big Data processing needs?										
BD 12	Does this organization has well trained IT Staff to deal with modern Big Data processing needs?										
BD 13	Please encircle data processing skills the IT professionals of this organization have?										
	<table border="0"> <tr> <td>A. NoSQL Databases Management Soft.</td> <td>F. SAS</td> </tr> <tr> <td>B. SQL Databases Management Soft.</td> <td>G. SPSS</td> </tr> <tr> <td>C. Hadoop (Pig/Hive)</td> <td>H. Java</td> </tr> <tr> <td>D. Spark</td> <td>I. Spread Sheet (Excel)</td> </tr> <tr> <td>E. R</td> <td>J. Others (Please Specify) _____</td> </tr> </table>		A. NoSQL Databases Management Soft.	F. SAS	B. SQL Databases Management Soft.	G. SPSS	C. Hadoop (Pig/Hive)	H. Java	D. Spark	I. Spread Sheet (Excel)	E. R
A. NoSQL Databases Management Soft.	F. SAS										
B. SQL Databases Management Soft.	G. SPSS										
C. Hadoop (Pig/Hive)	H. Java										
D. Spark	I. Spread Sheet (Excel)										
E. R	J. Others (Please Specify) _____										
BD 14	Please encircle statistical skills for data processing, IT professionals of this organization have?										
	<table border="0"> <tr> <td>A. Machine learning (Random forest, etc.)</td> <td>E. Decision Trees</td> </tr> <tr> <td>B. Supervised learning</td> <td>F. Data visualization methods</td> </tr> <tr> <td>C. Bayesian techniques</td> <td>G. Traditional statistical methods</td> </tr> <tr> <td>D. Neural networks</td> <td>H. Others (Please Specify) _____</td> </tr> </table>		A. Machine learning (Random forest, etc.)	E. Decision Trees	B. Supervised learning	F. Data visualization methods	C. Bayesian techniques	G. Traditional statistical methods	D. Neural networks	H. Others (Please Specify) _____	
A. Machine learning (Random forest, etc.)	E. Decision Trees										
B. Supervised learning	F. Data visualization methods										
C. Bayesian techniques	G. Traditional statistical methods										
D. Neural networks	H. Others (Please Specify) _____										
BD 15	Please encircle data processing software's are being used in this organization?										
	<table border="0"> <tr> <td>A. NoSQL Databases Management Soft.</td> <td>F. SAS</td> </tr> <tr> <td>B. SQL Databases Management Soft.</td> <td>G. SPSS</td> </tr> <tr> <td>C. Hadoop (Pig/Hive)</td> <td>H. Java</td> </tr> <tr> <td>D. Spark</td> <td>I. Spread Sheet (Excel)</td> </tr> <tr> <td>E. R</td> <td>J. Others (Please Specify) _____</td> </tr> </table>		A. NoSQL Databases Management Soft.	F. SAS	B. SQL Databases Management Soft.	G. SPSS	C. Hadoop (Pig/Hive)	H. Java	D. Spark	I. Spread Sheet (Excel)	E. R
A. NoSQL Databases Management Soft.	F. SAS										
B. SQL Databases Management Soft.	G. SPSS										
C. Hadoop (Pig/Hive)	H. Java										
D. Spark	I. Spread Sheet (Excel)										
E. R	J. Others (Please Specify) _____										
BD 16	Does this organization currently or in future have a plan to train IT Staff for data processing skills and techniques?										
BD 17	Please encircle data processing skills and techniques this organization plan to train its IT Staff?										
	<table border="0"> <tr> <td>A. NoSQL Databases Management Soft.</td> <td>F. SAS</td> </tr> <tr> <td>B. SQL Databases Management Soft.</td> <td>G. SPSS</td> </tr> <tr> <td>C. Hadoop (Pig/Hive)</td> <td>H. Java</td> </tr> <tr> <td>D. Spark</td> <td>I. Spread Sheet (Excel)</td> </tr> <tr> <td>E. R</td> <td>J. Others (Please Specify) _____</td> </tr> </table>		A. NoSQL Databases Management Soft.	F. SAS	B. SQL Databases Management Soft.	G. SPSS	C. Hadoop (Pig/Hive)	H. Java	D. Spark	I. Spread Sheet (Excel)	E. R
A. NoSQL Databases Management Soft.	F. SAS										
B. SQL Databases Management Soft.	G. SPSS										
C. Hadoop (Pig/Hive)	H. Java										
D. Spark	I. Spread Sheet (Excel)										
E. R	J. Others (Please Specify) _____										
BD 18	Does your organization hire some other IT firms for its data solution needs? (Like data warehousing, analytics, Cloud computing etc.)										
BD 19	If yes, then specify names of the Firms please. _____										
BD 20	Is there any provincial or federal government organization generates data which is in high interest of your organization?										
BD 21	If yes, then specify names of the Organizations please. _____										
BD 22	Is this organization in collaboration with other statistical or non-statistical government organization performs data recording/processing/management activities?										
BD 23	If yes, then specify names of the Organizations please. _____										

BD 24	There are several reasons due to which government organizations are not moving towards the adoption of Big Data for organizational use. Please rate the following reasons.						
	Reasons for Lacking Big Data Use		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
	A	Limited Awareness about Big Data	1	2	3	4	5
	B	Technological Stagnation	1	2	3	4	5
	C	Limited Resources	1	2	3	4	5
	D	Lack of Research & Research Environment	1	2	3	4	5
	E	Political Influences	1	2	3	4	5
	F	Legal Constraints	1	2	3	4	5
	G	Socio-Economic & Development Challenges	1	2	3	4	5
	H	Interests of the Higher Authority of the Organization	1	2	3	4	5
	I	Non-atomization in organizational work	1	2	3	4	5
	J	Lack of Statistically Advanced Analytical Skills	1	2	3	4	5
	K	Lack of Advanced Data Processing Skills	1	2	3	4	5
	L	Uncertainty about its usefulness	1	2	3	4	5
	M	Privacy & Policy Concerns	1	2	3	4	5
	N	Others (Please Specify)	1	2	3	4	5

P A R T - 3																
RATIONALIZATION OF STATISTICAL CADRE														RC		
RC 1	Is there any STATISTICAL/DATA PROCESSING post exist in your organization?										Yes.....1 » Continue No.....2 » Skip to <b>RC 3</b>					
Please rationalize Statistical Cadre in your organization in the table given below.																
RC 2	<b>Name of Statistical / Data Processing Post</b> <small>(Statistic Reporter, Statistical Assistant, Statistical Officer, Data Analyst, Data Programmer, Assistant Director Stats., Director Statistics, Statistician, Chief Statistician Etc.)</small>			<b>BS</b>	<b>No. of posts</b>		<b>Functions / Activities</b> (Please encircle the relevant)									
	<b>Sanctioned</b>	<b>Vacant</b>	Data Collection		Editing/Coding	Tabulation	Data Entry	Data Management	Data Analysis	Report Writing	Data Dissemination	Survey Planning / Survey Methodology	Others (Please Specify)			
				C	E	T	P	M	A	R	D	S				
				C	E	T	P	M	A	R	D	S				
				C	E	T	P	M	A	R	D	S				
				C	E	T	P	M	A	R	D	S				
				C	E	T	P	M	A	R	D	S				
				C	E	T	P	M	A	R	D	S				
				C	E	T	P	M	A	R	D	S				
RC 3	We are thankful for your cooperation in providing the information about the production of Official Statistics in your organization. <i>Please give your valuable comments about the use of Big Data in Official Statistics or any else.</i>															

**THANKS**

Note: In case of any query, please feel free to contact the focal person during office hours.

**Syed Wasim Abbas**  
 Focal Person (SOS-Pak)  
 +92) 333 6550630  
 alsyed\_edu@hotmail.com

Scan to download a copy of this Questionnaire...

or

Email us at: alsyed\_edu@hotmail.com

