TRAININGS SERIES

Surveys | Statistics | Big Data

By: Juma AlHosani

Contents

1 OVERVIEW

Duration: On average each session will be 02+ hours long
Mode: Interactive, learner-centered sessions
Format: Readings, presentations, quizzes, group projects, practical experience, Assignment, case studies based training sessions.
Resources: A basic Learning Management System [e.g. Moodle] to manage participants' and course related activities and data.

2. OBJECTIVES

Objectives of this course are to introduce participants to the skills and resources needed to design, conduct and analyze a survey. Participants will learn:

1. How to design surveys
2. How to sample in surveys
3. How to develop, evaluate, and ask survey questions
4. How to measure survey reliability and validity
5. How to conduct self-administered and mail surveys
6. How to decrease survey nonresponse
7. How to reduce error in survey research
8. How to do post-collection survey data processing
9. How to conduct survey research with integrity
10. What are latest survey data analysis tools
11. What is big data and how to benefit from it.

3 SURVEY TOPICS

3.1 Month 1 - SURVEY DESIGN

	SURVEY DESIGN - Importance of Survey Design - Why do a survey? - What are the goals? - Sample versus a Census - Defining the population precisely - Survey Modes (Telephone, Mail, Web, Face-to-face) - What are some typical uses of surveys? - Case Studies

3.2 Month 2 - SAMPLING

	METHODS AND THEORY OF SAMPLE DESIGN - Statistical concepts and techniques in sample design - Simple random sampling - Stratification - Systematic sampling - Cluster and multi-stage sampling - Use of technology for sampling - Sampling aspect of web surveys - Web surveys vs Internet and online surveys

3.3 Month 3 – DEVELOPING QUESTIONS

	DEVELOPING QUESTIONS - Reliability and validity - Tips /checklist for writing good questions - Design of self-administered questionnaires - Wording of questions (strategies for factual and non-factual questions) - Cognitive aspects - Order of response alternatives - Open versus closed questions - Handling sensitive topics - Combining individual questions into a meaningful questionnaire - Questions and levels of measurement • Categorical (nominal) questions • Ordinal questions • Interval questions • Open-ended questions

Related topics – Month 3

COGNITION, COMMUNICATION, AND SURVEY MEASUREMENT

Topics will be covered from cognitive and social psychology pertaining to issues such as language comprehension, information storage and retrieval, autobiographical memory, social judgment, and the communicative dynamics of survey interviewing, to understand how respondents deal with the questions asked and how they arrive at an answer.

3.4 Month 4 – DATA COLLECTION

	DATA COLLECTION METHODS - Online, Email, Face-to-face, Telephone, Mail - Response rate - Increasing response rate in web-based surveys - Response rate and reward - Building your own online panel via email and online digital media - Follow up procedures

Related topics – Month 4

Ensuring survey engagement, survey gamification, questionnaire length, survey satisficing
Avoiding biases, impact of biases, ways to avoid biases

3.5 Month 5 – DATA ANALYSIS

	DATA ANALYSIS METHODS - Data cleaning and file preparation - Classification systems and recodes - Descriptive statistics and hypothesis testing - Analysis of variance - Data reduction through factor and/or cluster analysis and the development of indices

	- Cross-classification of categorical data and the measurement of association Multivariate linear regression tools - Dummy-variable regression and multiple classification analysis - Causal analysis and multiple dependent variables - Measurement errors and statistical analysis - Report writing - Graphics and the presentation of data

3.6 Month 6 – DEALING WITH ERRORS

	DEALING WITH ERRORS - Response & non-response errors - Unit non-response due to refusal - Non-response and measurement error - Why people agree to participate in surveys - Demographic profile of survey respondents and its implications.

3.7 Month 7 – APPLICATIONS of WEB SURVEYS

	APPLICATIONS OF THE [web] SURVEYS - Employee surveys as catalysts for change: Turning Data into Action - Ethical issues in survey research - Customer satisfaction surveys - General population surveys - Business surveys - Online panels - Online community surveys - Entertainment polls, questionnaires and quizzes - Psychological web experiments

3.8 Month 8 – SURVEYS & STATISTICS

	SURVEYS & STATISTICS - What is statistics - Descriptive, parametric and non-parametric statistics - Hypothesis and its kinds **Descriptive Statistics:** - Mean, Median, Mode, Standard Deviation, Correlations, Confidence Interval **Parametric** - Pearson r correlation, Linear/multiple regression **Non-Parametric** - Testing categorical data, chi square, logistic regression

3.9 Month 9 – SURVEY ANALYSIS in R

	SURVEY ANALYSIS IN R

	- Means, totals, ratios, quantiles, contingency tables, regression models, loglinear models, survival curves, rank tests, for the whole sample and for domains. - Multistage sampling with or without replacement. - Graphics - Support for using multiply imputed data - Multivariate analysis: principal components, factor analysis (experimental). - Likelihood ratio (Rao-Scott) tests for glms, Cox models, loglinear models.

3.10 Month 10 – INTRODUCTION TO BIG DATA

	INTRODUCTION TO BIG DATA **Defining Big Data** - The four dimensions of Big Data: volume, velocity, variety, veracity - Introducing the Storage, MapReduce and Query Stack **Delivering business benefit from Big Data** - Establishing the business importance of Big Data - Addressing the challenge of extracting useful data - Integrating Big Data with traditional data

3.11 Month 11 – BIG DATA STORES

	OVERVIEW OF BIG DATA STORES
	- Data models: key value, graph, document, column–family
	  - Hadoop Distributed File System,
	  - HBase, Hive,
	  - Cassandra,
	  - Hypertable,
	  - Amazon S3,
	  - BigTable,
	  - DynamoDB,
	  - MongoDB,
	  - Redis,
	  - Riak,
	  - Neo4J

3.12 Month 12 – BIG DATA STRATEGY

	DEVELOPING A BIG DATA STRATEGY **Defining a Big Data strategy for your organization** - Establishing your Big Data needs - Meeting business goals with timely data - Evaluating commercial Big Data tools - Managing organizational expectations **Enabling analytic innovation** - Focusing on business importance - Selecting the correct tools - Achieving timely results - Establishing the business importance of Big Data

	- Addressing the challenge of extracting useful data - Integrating Big Data with traditional data **Implementing a Big Data Solution** - Selecting suitable vendors and hosting options - Balancing costs against business value - Keeping ahead of the curve

5 Innovation During Presentations

5.1 Real-time feedback during the presentations.

Participants can use their mobile devices to provide their feedback.
PollEverywhere can be used initially.
https://www.polleverywhere.com/

5.2 Infographics for Presentations

https://www.visme.co/

5.3 Interactive Presentations

During the presentation, presenter or participants can use touch screen panels to deliver presentation.


Build on Windows
Play on Any Touch Device
Walls
Kiosks
Tablets
- Nobu Matsuhisa
Play Your Experiences On
Windows
iPad
android
chrome
TIZEN

6 Samples of Session Resources

6.1 Day Plan of A 2-Hour Session

Date	Session	Topic & Activities	Required Readings	Material	Assignment
01st August 2017	1 of 2 Session 2:	Introduction to survey methodology & the tailored design method Brainstorming, group discussion,	Groves et al. (2009) – Ch. 1 & 2 Dillman – Ch. 1 Miller (2012) Squire (1988)	- Group activity -	Write 5 distinct points which describe survey methodology

6.2 Sample Assignment Project

Participants will be divide groups of two or three according to their research interests. The groups then will need to decide on a survey research project that addresses a particular problem and answers a specific research question(s). Participants will not implement the survey for this project, but each group will develop a short questionnaire and write a research proposal describing the survey.

The research proposal will be approximately 20-25 double-spaced pages (12 pt. font, formatted using APA, 5th edition) not including the appendices or references.

The proposal should include the following:

1. Introduction

In this section, participants will delineate the research problem, describe the purpose of your survey, provide a brief review of the literature, and list the research question(s) your project proposes to answer.

2. Method

a. Describe the survey and the constructs to measure.

b. Describe:

i. Target population – who to study?

ii. Sampling frame – how to identify the people who have a chance to be included in the survey?

iii. Sample design – how to select members of the sample, and how many will be selected?

iv. Survey mode – how to contact members of the sample, how to ask the questions and collect answers, and how much effort will be devoted to collecting data from those reluctant to respond?

c. Explain how to evaluate the instrument prior to data collection.

d. Note any ethical issues and your approach to dealing with them.

3. Post collection processing of data

Briefly describe the plan for, how to enter, code, and check the data after collection.

4. Limitations

No survey is free of error. Describe issues of error that your design is not able to overcome or areas where you expect error will be introduced.

5. Cover letter/email/interviewer protocol

Include a copy of the communication that will accompany your survey as an appendix. This should be no longer than one page; it can be single-spaced, 12 pt. font.

6. Survey

Include a copy of your instrument as an appendix. This should be no longer than four pages; it can be single-spaced, 12 pt. font.