

PROJECT PROPOSAL
IDS 570 Statistics for Management

Amol Modi, Arjit Jain, Chandana Dibadahalli Kumar, Madhav Mehta, Parshv Sanghvi &
Siddhant Yadav

Provide a detailed description to the following prompts (50% each)

1. Identify a data set relevant for the statistical analyses that your team expects to use for the Final Project.

- a. How have you obtained the data? For example, a workplace of one of the team members; an online source (provide a link).

The data was obtained from Prof. Zack Kertcher.

- b. How were the data collected?

The data has been collect by surveying federal government employees in 2010.

- c. What are the main characteristics of the data?

Number of observations: 263,475

Number of variables: 21

Sr. No	Variable Name	Variable Type	Description
1	AGENCY	FACTOR	Name of the Government Agency
2	Q3	INT	I feel encouraged to come up with new and better ways of doing things.
3	Q4	INT	My work gives me a feeling of personal accomplishment.
4	Q5	INT	I like the kind of work I do.
5	Q6	INT	I know what is expected of me on the job.
6	Q11	INT	My talents are used well in the workplace.
7	Q13	INT	The work I do is important.
8	Q47	INT	Supervisors/team leaders in my work unit support employee development.

9	Q48	INT	My supervisor/team leader listens to what I have to say.
10	Q53	INT	In my organization, leaders generate high levels of motivation and commitment in the workforce.
11	Q56	INT	Managers communicate the goals and priorities of the organization.
12	Q63	INT	How satisfied are you with your involvement in decisions that affect your work?
13	Q67	INT	How satisfied are you with your opportunity to get a better job in your organization?
14	Q69	INT	Considering everything, how satisfied are you with your job?
15	Q70	INT	Considering everything, how satisfied are you with your pay?
16	DLOC	FACTOR	Where do you work?
17	DSUPER	FACTOR	What is your supervisory status?
18	DSEX	FACTOR	What is your sex?
19	DAGEGRP	FACTOR	What is your age group?
20	Employee Engagement Index	INT	The aggregation of responses for variables Q3, Q4, Q6, Q11, Q47, Q48, Q53 and Q56
21	Job Satisfaction Index	INT	The aggregation of response for variables Q4, Q5, Q13, Q63, Q67, Q69 and Q70

2. Elaborate the following:

- a. What is your research question? Refer to lecture slides
 - i. Is there a relationship between Job Satisfaction and Employee Engagement Index?
 - ii. Does the Employee Engagement Index depend between Location, Supervisory Status, Sex and Age Group?

- b. What are the dependent, independent and control variables? Refer to lecture slides

Independent Variable	Q3, Q4, Q5, Q6, Q11, Q13, Q47, Q48, Q53, Q56, Q63, Q67, Q69 and Q70
Dependent Variable	Employee Engagement Index and Job Satisfaction Index
Control Variable	AGENCY, DLOC, DSUPER, DSEX and DAGEGRP

- c. What are your hypotheses? Refer to lecture slides
- There is a relationship between Job Satisfaction and Employee Engagement Index (EEI).
 - There is no difference in the average EEI between Males and Females.
 - There is no difference in the average EEI between Age Groups.
 - There is no difference in the average EEI between employee located in Headquarters and Field.
- d. What challenges do you anticipate experiencing during this project? For example, the data seem “messy.”
- There are a lot missing values in various variables. It would be a huge challenge to remove the rows with the missing value in the variables.
 - There are certain variables which have values equal to ‘X’. Because of this the variable is being treated as Factor. We would be required to replace values equal to ‘X’ to NA and then remove the rows with the missing value in the variables.