Example email. Much modified from work by Anita Lok, used with her permission – Caroline Moore-Kochlacs, Nov 2019

A.B. Clark
ALY 6100, Data-Driven Decision Making
Week 2 Assignment: Data Collection

From: abc@huskey.neu.edu
To: c.moore-kochlacs@northeastern.edu
Subject: Project Oxygen: Data Collection Approach

Caroline,

The data science team is ready to get started on our areas of Project Oxygen. As you requested, I've put together plan for the first two weeks of data collection.

The question at the heart of Project Oxygen is: "Do managers matter?" (Gavin, 2013). Because so many Google employees are skeptical that managers are valuable, our first phase of analysis will be an attempt to prove that managers do not matter (Gavin, 2013).[1] If our research disproves that managers do not matter, and that managers *do* matter, our goal is to also identify granular, specific, and measurable management traits that are valuable.

I want to start by proposing some scope parameter for the first phase of our Project Oxygen analysis. I suggest we look at:

- records from the past 2 years, 2007-9
- managers working in 3 departments
- managers working out of the Mountain View Google Campus

I believe this will be a sufficient, but not overwhelming amount of data to start analyzing. I'd ask your help in determining which 3 departments we should start with, as I will need your help interfacing with those departments. Finally, by restricting the data we're analyzing to managers working at the Mountain View campus, we will have the ability to have in person interviews with managers and team members, if the need arises.

I propose, for the time being, we use the following as potential indicators of manager's value, negative or positive, to their direct reports and team members:

1. Managers leading to employees leaving Google. We will measure this by counting and reading mentions of management issues in exit-interview data and transcripts (Gavin, 2013).

[1] Proving by disproving a hypothesis is a common strategy in mathematical proofs

2. Manager performance on annual and semi-annual reviews (Gavin, 2013). We specifically want to look at how individual manager's direct reports and own managers rated them in terms of performance and satisfaction (Gavin, 2013).
3. Team turnover (Gavin, 2013). Employees leaving teams, either for other positions internally or externally, may indicate a manager's value. Lateral moves internally and leaving Google may indicate negative value. Internal promotions may indicate positive value.
4. Team performance. In departments in which team performance metrics are available (e.g. performance on sales targets or launch goals), progress relative to team objectives may indicate negative or positive value of managers.

I propose we begin our data collection by gathering and collecting the following data. In all cases, the data should be collected for:

- Years: 2007-2009
- Mountain view campus managers, their managers, and their direct reports. (Their managers and their direct reports, independent of geographic location.)
- Departments: Ideally this data will be collected for all departments, to cover anyone who moved between departments and have the data for later phases of analysis. If it's significantly faster or easier to collect for a subset of departments, just the departments we're analyzing in the first phase

Google management/employee hierarchy and position data.

- Data segmentation: Monthly basis, first day of the month, 2007-9.
- Data fields
    - Employee Name and/or ID
    - Position
    - Department
    - Team name, if exists
    - Direct manager's employee ID
- Why? This data will allow us to identify which teams and employees each Google manager was/are responsible for. This will allow us to related much of the rest of the data to specific managers and give us a detailed picture of team history and turnover, in combination with the exit interview data.
- Collection: HR should have this data, but I am not sure whether it is already structured in this way. Some manual or organizing work may need to go into making the data HR can give us useable.

Exit interview data about managers. Specifically:

- Data fields
    - Exit interview date
    - Leaving employee's name or ID
    - Any survey questions about managers and their answers
    - Any mention of managers in text transcripts

- Why?
    - This will give us data about how leaving employees perceived their managers. An employee leaving due to their manager would be, in most circumstances, an indicator of a manager's negative value.
- Collection
    - HR should have this data, but again, I am now sure how structured it is / how much work it will take to get it organized. For the text transcripts, I suggest we search all transcript text for the keywords "manager" and the names of the employees managers during their time at Google. Only the sections of text around those keywords should be extracted.

Manager performance reviews. Complete annual and semi-annual performance review data. Our analysis will start with by looking at their performance and satisfaction ratings and comments from their direct reports and own managers. In aggregate, we expect a manager's performance ratings to be an indication of their value. (We will want to cross-validate this with other indicators.) In later phases of the analysis, we will look at other sections of the review to see if specific traits which are assessed in the reviews are associated with manager value. This data will come from HR.

Team performance metrics. For teams in which team performance can be explicitly measured, e.g. Sales performance on goals, we would like complete performance data. The details of this will be clearer when we know which 3 departments we'll focus on to start.

Please let me know when you can meet to discuss this plan. In addition to determining which 3 departments we'll start analysis on, I'd like to discuss whether we should use employee name or ID in our analyses.

Once approved, the team will get started on the data pulls. I am hopeful all the data can be collected in the next two weeks, though we're going to need to interface with HR to get a better estimate.

Thanks, A.B.

References

Bartlett, R. (2013). A practitioner's guide to business analytics: Using data analysis tools to improve your organizations decision making and strategy. New York: McGraw-Hill.

Garvin, D. A. (2013, December). How Google Sold Its Engineers on Management. Retrieved July 14, 2018, from https://hbr.org/2013/12/how-google-sold-its-engineers-on-management