# # Analytic Exercise for Evidence Based Practice Training

## How to run this

\*\*Option 1:\*\* Give this data to individuals or small groups and ask them to compile a written report. I would encourage people to post their report publicly and open it up for critique from the analytics community.

\*\*Option 2:\*\* Give this data to individuals or small groups. After a certain time period allowed for analytics, have the group present their findings to a panel of three role-played executives, each with an ingoing set of beliefs about the situation. Please <a href="mailto:keith.mcnulty@gmail.com">email me</a> for role play scripts if you choose this option.

# ## The context

You are working for a North American sales company that has offices in six locations. Recently the CEO of this company has become interested in the performance of its managers. She is interested in knowing what **measurable factors influence whether or not a manager performs well or performs poorly.**

# ## Data available

The data available includes the following fields for 571 managers:

\* `employee\_id` for each manager

\* `performance\_group` of each manager: Bottom performer, Middle performer, Top performer

\* `yrs\_employed`: Total length of time employed in years

\* `manager\_hire`: whether or not the individual was hired directly to be a manager (Y) or promoted to manager (N)

\* `test\_score`: score on a test given to all managers

\* `group\_size`: the number of employees in the group they are responsible for

\* `concern\_flag`: whether or not the individual has been the subject of a complaint by a member of their group

\* `mobile\_flag`: whether or not the individual works mobile (Y) or in the office (N)

\* `customers`: the number of customer accounts the manager is responsible for

\* `high\_hours\_flag`: whether or not the manager has entered unusually high hours into their timesheet in the past year

\* `transfers`: the number of transfer requests coming from the manager's group while they have been a manager

\* `reduced\_schedule`: whether the manager works part time (Y) or full time (N)

\* `city`: the current office of the manager.

The data is available in an RData file for use in R (which ensures the data types are correct), or a csv file for use in any other analytics tool, where you may need to convert to the correct data types. Choose whichever you prefer to work with.

# ## Your task

Using whatever method you deem appropriate, try to determine:

1. Which variables have an effect on manager performance?

2. For those variables which have an effect, how 'strong' is the effect?

3. What might be an explanation for such effects?

4. What are the limitations of your analysis?

The purpose of this exercise is to expose the typical challenges of communicating analytical results to a non-expert audience. Be prepared to present and explain your results to an audience of mixed experience in analytics, whether via a written report or a verbal presentation/discussion.

\*\*Do not try to use an automated analytics platform for this work\*\*. These platforms are very poor substitutes for strong methodology and nuanced interpretation. At best they produce limited insight and at worst they produce results that are patently wrong. Remember that this exercise has been created to help build your portfolio of analytics skills and your ability to make nuanced interpretations and communicate them to others. Delegating this to a one size fits all analytics platform will not help.

Enjoy!