# Presentation of Findings

## Executive Summary and Implications

### A. Develop an executive summary using your data and results from task 2. The summary should be written for a *technical audience* of your data-analytics peers or members of your team and should include *each* of the following:

- a statement of the problem and the hypothesis
- a summary of the data-analysis process
- an outline of the findings
- an explanation of the limitations of the techniques and tools used
- a summary of proposed actions
- expected benefits of the study (be as specific and quantitative as possible)

**Problem Statement and Hypothesis:**

Predictive analytics was used to forecast energy capacity and supply mismatches in Puerto Rico's energy grid under stress conditions. The hypothesis declared that analyzing time series data and grid performance capacity metrics could improve forecasting accuracy. This in turn would facilitate both planning and emergency response for energy supply and capacity mismatches in high-stress events.

**Summary of the Data Analysis Process:**

Two datasets were used in this analysis: one for predicted hourly energy generation for various technologies over a one-year period and another for energy generation capacities under different scenarios. The Pandas library was used for data extraction, preparation, and manipulation. Unnecessary columns were dropped, column names were standardized across the two dataframes, the data was resampled to a daily frequency to reduce computational load, calculations were made to obtain the energy supply and capacity mismatch for each energy type, then the two dataframes were merged on technology type. This new dataset contained daily energy and supply mismatch data for 9 technology types.

**Key Findings:**

Significant mismatches were identified for technology types like Utility Photovoltaics, Distributed Photovoltaics, and Diesel. The ARIMA model effectively forecasted future mismatches, providing information that can be used to plan infrastructure development and energy resource adjustments.

**Limitations of the Techniques and Tools Used:**

Both the Pandas library and the ARIMA model can be resource-intensive with very large datasets, which could limit the scalability of the model used in this analysis. The analysis was also constrained by the scope of the historical data, which only covered one year and was based on predictive models rather than real-world observations.

**Proposed Actions:**

Future infrastructure planning should consider the results of this analysis in effectively distributing generated energy and reducing mismatches. Commensurate battery storage and other responsive energy resources can be used to smooth out forecasted mismatches to ensure a more reliable energy supply.

**Expected Benefits:**

The reliability of Puerto Rico's power grid during stress conditions should improve, reducing energy shortages and wasted energy generation. Proactive infrastructure upgrades and the integration of battery storage could lead to a 10-20% reduction in operational costs associated with emergency energy distribution.