Intel Data Center

INTRODUCTION: Intel, the semiconductor manufacturing powerhouse, is planning on building a new data center. Energy availability and usage are some of the key considerations in deciding on a location of the data center. For example, which regions produce a surplus of energy, and are therefore more likely to provide energy at cheaper prices? Which regions rely more on renewable energy sources?

In this project, co-designed with Intel's Sustainability Team, you'll write SQL queries that will power your analysis and create visualizations that will help the Intel team select the best location for the new data center.

— Data Sets

`intel.energy_data`: Contains information about daily energy production and consumption for different regions in the United States.

- `balancing_authority` - A Balancing Authority is responsible for maintaining the electricity balance within its region. This is a company that makes sure electricity is being exchanged between electric providers and regions so that no region runs out of electricity due to high demand.
- `date` - The date the energy was produced.
- `region` - The electric service area within a geographic area of the USA. e.g. California, Midwest, etc.

- `time_at_end_of_hour` - The time and date after energy was generated, .e.g., energy generated between 1pm-2pm will show up as 2pm in this field.
- `demand` - The energy demand in megawatts (MW) on the grid (what the houses/business are using).
- `net_generation` - The energy produced in MW in the region by all sources e.g., wind, coal, nuclear, etc.
- `all_petroleum_products` - The energy produced in MW by petroleum products.
- `coal` - The energy produced in MW by all coal products
- `hydropower_and_pumped_storage` - The energy produced in MW by water power and pumped heat sources.
- `natural_gas` - The energy produced in MW by natural gas sources
- `nuclear` - The energy produced in MW from nuclear fuel sources
- `solar` - The energy produced in MW by solar panels and other solar energy capturing methods.
- `wind` - The energy produced in MW from wind turbines and other wind sources.

`intel.power_plants`: Contains general information about power plants in the United States.

- `plant_name` - The name of the power plant.
- `plant_code` - The unique identifier of the plant.
- `region` - The region in the US where the power plant is located. Matches the regions in the intel.energy_data
- `state` - The state where the power plant is located.
- `primary_technology` - The primary technology used to generate electricity at the power plant.

`intel.energy_by_plant`: Contains total energy production information at the plant for the year 2022.

- `plant_name` - The name of the power plant.
- `plant_code` - The unique identifier of the plant.
- `energy_type` - The kind of energy generated by the power plant. Either renewable energy or fossil fuel.
- `energy_generated_mw` - The total energy generated, in MegaWatts, at the plant for the year 2022.

A. Use the built-in visualizer in the SQL app to plot a line graph of the energy generated for each hour of the day and colored by the region. If done correctly you should have two lines in your visualization.

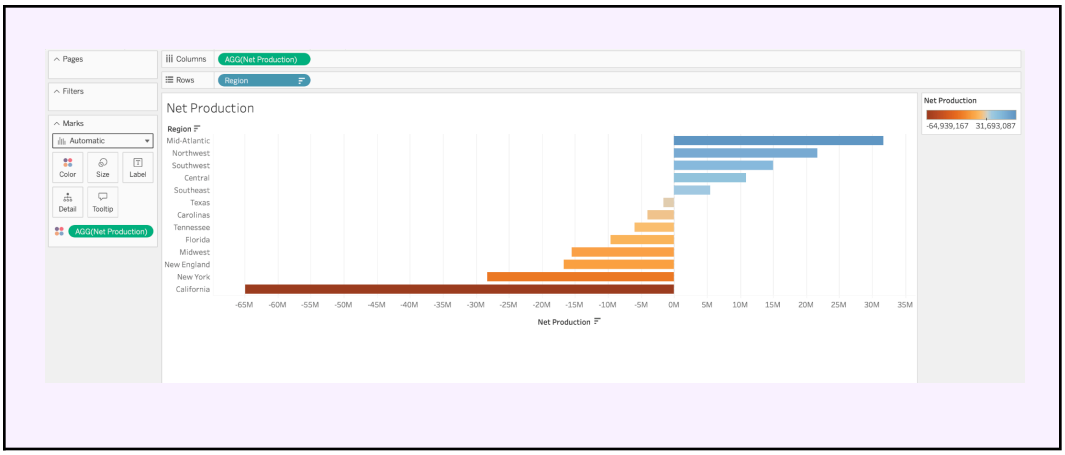
region	hour_of_day	total_renewable_energy
California	0	1706137
California	1	1580759
California	2	1492731
California	3	1445163
California	4	1411726
California	5	1434938
California	6	1535721

B. What can you say about the renewable energy generation between California (CAL) and the Pacific Northwest (NW)?

Overall, the Northwest produces much more energy than California does. Both regions dip to their lowest production point at hour 4, and the hours of highest production between hours 9–20 for the Northwest and between 9–16 for California. There is a small dip in the center of these high points, meaning production was very high, went down just a bit, and then rose once again before falling off for the day.

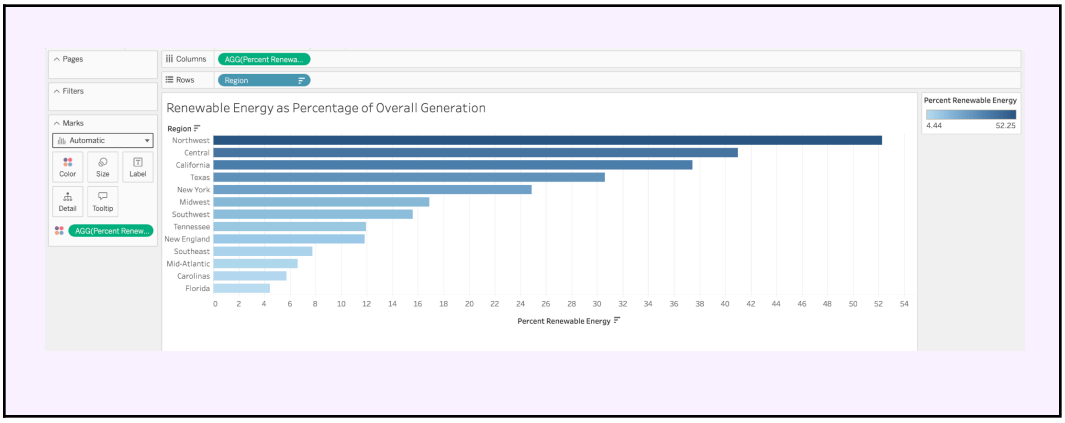
– **Task 4:** Visualizing and Analyzing Using Tableau

A. On the “Net Production” sheet, create a bar chart of net production , by region. Sort the chart in *descending* order, from tallest to smallest.

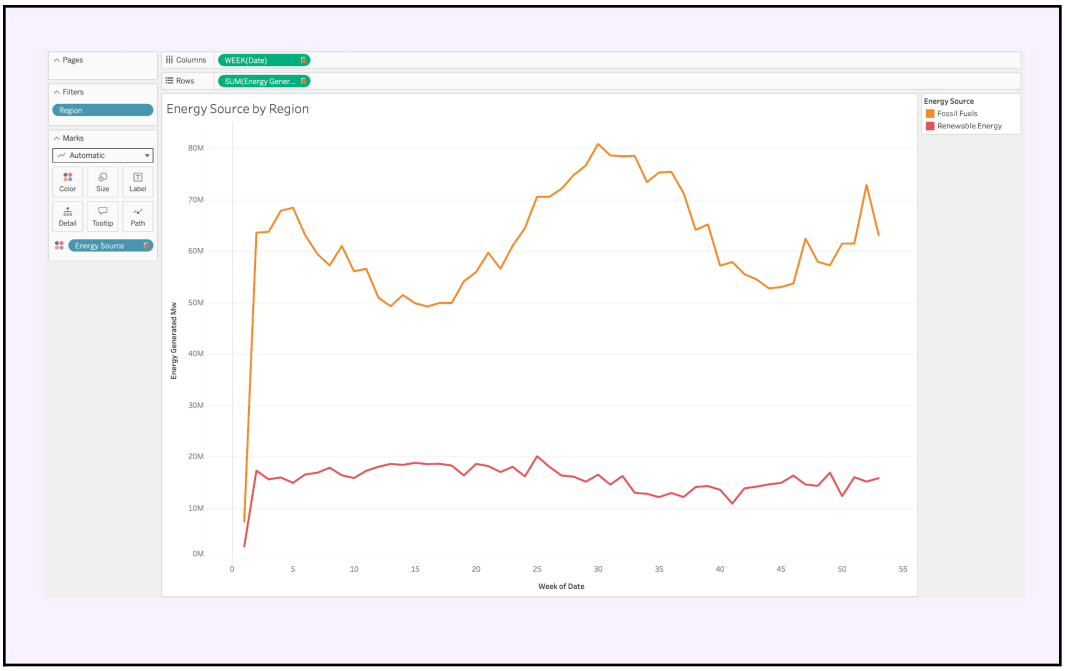


B. Next, on the “Renewable Energy” sheet, create a bar chart illustrating which regions generate the greatest percentage of renewable energy.

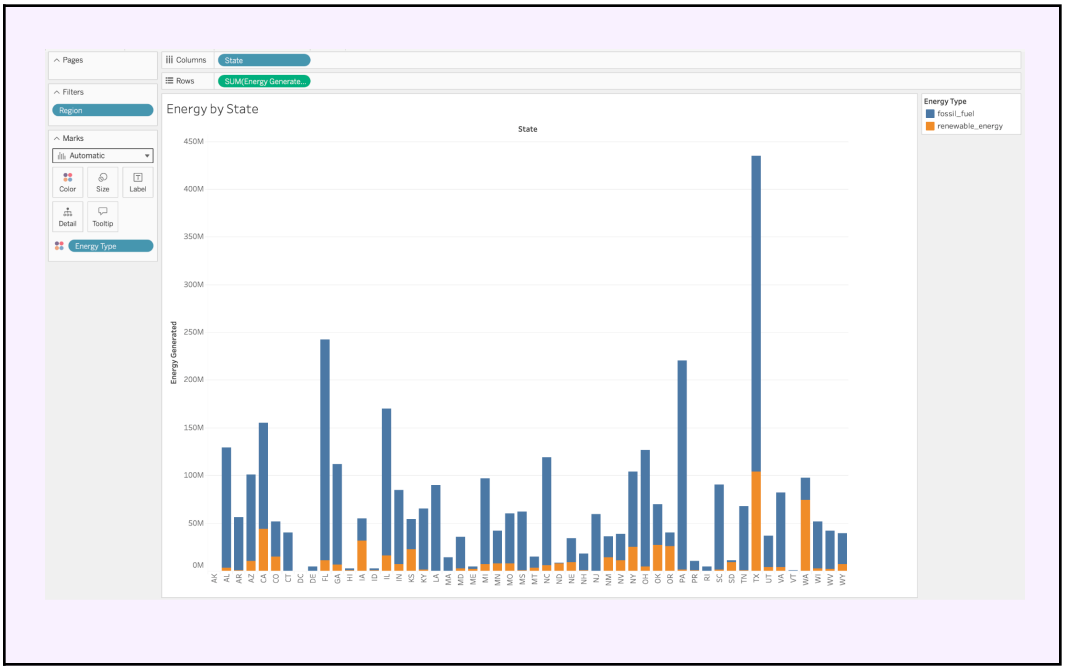
   Create a bar chart in descending order of regions with the most renewable energy percentage.



C. On the “Energy Source by Region” sheet, create a line chart of the energy generated for each energy source (fossil fuels & renewable energy) at the weekly date level. Add a filter for the region to your chart.



D. On the "Energy by State" sheet create a bar chart of the total energy generated by each state and energy type. Color the bars by energy type. Include a region filter in your chart to reduce the amount of bars shown.



– **Task 5:** Communicating Results

Your manager wants you to share the visualizations you created in Task 3 with the Sustainability team for visibility. She has created a dashboard with your visualizations (see the "Dashboard" sheet in Tableau) and has asked you to write a short paragraph explaining which region you recommend that the next data center be built.

A. What **region** and **state** do you think is best and why?

Looking at the data, the regions with the highest surplus of total produced energy are the Mid-Atlantic, Northwest, and Southwest, which first points us towards ruling out the Southeast and Midwest sections of the USA for the next data center. Narrowing in on these three regions, the Northwest has the highest production of renewable energy overall, whereas the Southwest is placed at 7th and the Mid-Atlantic region is placed at 11th for overall percentage generation.

Now that we have a region that has a surplus of energy produced (which makes energy cheaper) and produces a large percentage overall of renewable energy (which makes energy cleaner), it's time to look closer at the states in that region to determine which is the best contender for the data center. The states in the Northwest are Washington, Oregon, and Idaho. The state producing the most energy and most renewable energy out of the three is Washington, making it the best contender. If we look at another region from the first part that has a surplus of energy but may fall back shorter in the percentage generated portion, we have the Southwest region (Nevada, Utah, Colorado, Arizona, and New Mexico) I feel the best option among these would be Colorado based on the ratio of fossil fuel to renewable energy use and amount of energy produced.