



STATE OF CALIFORNIA

DUTY STATEMENT

CEC-004 (Revised 7/2020)

Classification(s): Research Data Specialist II

Working Title: Data Pipeline Developer

Position Number: 535-153-5758-XXX

Division/Office: Energy Data and Analytics Office

Collective Bargaining Identifier (CBID): R01

Work Week Group (WWG): E

Effective Date: March 1, 2021

Conflict of Interest (COI): Yes No

Job Description

The Energy Data and Analytics Office is responsible for organization-wide infrastructure, tools, automation, security, and governance required to effectively manage and analyze the Commission's data assets.

Under the general direction of the Assistant Executive Director for Energy Data and Analytics, the incumbent is responsible for developing and maintaining the most complex data pipelines, relational data models, and data views to efficiently ingest, process, validate, format, store, and query structured, semi-structure, and unstructured data. With minimal technical oversight, the incumbent independently accomplishes the above, along with other related data engineering tasks, using cloud computing resources, data integration and business intelligence tools, and custom code. Additionally, the incumbent effectively communicates and coordinates technical requirements of data pipelines and relational data models with other staff supporting pipeline development and with non-technical data end users and project managers. The incumbent examines and manipulate data using an object-oriented programming language, designs and implements relational table structures and data views, and queries relational data using Structured Query Language (SQL). When necessary, the incumbent can learn and become proficient with new data management environments and data integration tools. The incumbent will develop recommendations for potential improvements to products or processes that support the overall analytical needs of the Energy Commission.

Essential Duties

45% Serve as the prime resource for developing and maintaining assigned data pipelines for data the Energy Commission regularly receives via regulatory requirements, interagency data sharing agreements, and public data sources. Utilize data integration software and custom

code for more complex data validation, formatting, and organization tasks. Work with support staff, data end users, and program managers to understand technical requirements. Key responsibilities include: developing and maintaining extract, transform, and load (ETL) jobs using a data integration tool; writing or customizing additional ETL code as necessary; and automating data validation and quality checks.

30% Lead the development of custom data views and reports for the data warehouse. Work with program areas in developing more complex SQL queries for the data warehouse. Respond to ad-hoc requests for data queries and reports. Responsible for querying and manipulating data using SQL and querying and joining data across multiple datasets.

15% In coordination with data end users, develop relational data models for data stored in the Energy Commission's data warehouse. Responsible for ensuring data tables in the data warehouse are appropriately designed to satisfy analytical and reporting use cases for that data.

Marginal Duties

5% Advise staff and management on technical issues related to the implementation of new data collection efforts or the improvement of existing efforts.

5% Perform other duties as required, consistent with the specifications of the classification.

Working Conditions

Due to the COVID-19 pandemic, the majority of the duties will be performed in a telework environment, but may return to an office environment in the future. Regular and consistent attendance is essential to the successful performance in this position. Occasional travel may be required. Additional hours beyond an eight-hour workday or forty-hour workweek may be required.

Employee's Acknowledgement: I certify that I am able to perform, with or without the assistance of a reasonable accommodation, the essential duties of this position.

Employee's Name (Print): _____

Employee's Signature: _____ **Date:** _____

Supervisor's Acknowledgment: I certify this duty statement represents a current and accurate description of the essential functions of this position. I have discussed the duties of this position with and provided the above-named employee a copy of this duty statement.

Supervisor's Name (Print): Jason Harville

Supervisor's Signature: _____ **Date:** _____



STATE OF CALIFORNIA

DUTY STATEMENT

CEC-004 (Revised 7/2020)

Classification(s): Research Data Specialist I

Working Title: Associate Data Pipeline Developer

Position Number: 535-153-5742-XXX

Division/Office: Energy Data and Analytics Office

Collective Bargaining Identifier (CBID): R01

Work Week Group (WWG): 2

Effective Date: March 1, 2021

Conflict of Interest (COI): Yes No

Job Description

The Energy Data and Analytics Office is responsible for organization-wide infrastructure, tools, automation, security, and governance required to effectively manage and analyze the Commission's data assets.

Under the general direction of the Assistant Executive Director for Energy Data and Analytics, the incumbent is responsible for independently developing and maintaining data pipelines, relational data models, and data views to efficiently ingest, process, validate, format, store, and query structured, semi-structure, and unstructured data. Both independently and with support from other data pipeline staff as needed, the incumbent accomplishes the above, along with other related data engineering tasks, using cloud computing resources, data integration and business intelligence tools, and custom code. Additionally, the incumbent effectively communicates and coordinates technical requirements of data pipelines and relational data models with other staff supporting pipeline development and with non-technical data end users and project managers. The incumbent examines and manipulate data using an object-oriented programming language, designs and implements relational table structures and data views, and queries relational data using Structured Query Language (SQL). When necessary, the incumbent can learn and become proficient with new data management environments and data integration tools. The incumbent will develop recommendations for potential improvements to products or processes that support the overall analytical needs of the Energy Commission.

Essential Duties

45% Develop and maintain assigned data pipelines for data the Energy Commission regularly receives via regulatory requirements, interagency data sharing agreements, and public data sources. Develop, implement, and monitor systems and procedures for structuring and

validating data. Work with support staff, data end users, and program managers to understand technical requirements. Key responsibilities include: developing and maintaining extract, transform, and load (ETL) jobs using a data integration tool; writing or customizing additional ETL code as necessary; and automating data validation and quality checks.

30% Assist in the development of custom data views and reports for the data warehouse. Work with program areas to develop complex SQL queries for the data warehouse and assemble additional data as needed. Responsible for querying and manipulating data using SQL and querying and joining data across multiple datasets.

20% In coordination with data end users, develop relational data models for data stored in the Energy Commission's data warehouse. Responsible for ensuring data tables in the data warehouse are appropriately designed to satisfy analytical and reporting use cases for that data.

Marginal Duties

5% Perform other duties as required, consistent with the specifications of the classification.

Working Conditions

Due to the COVID-19 pandemic, the majority of the duties will be performed in a telework environment, but may return to an office environment in the future. Regular and consistent attendance is essential to the successful performance in this position. Occasional travel may be required. Additional hours beyond an eight-hour workday or forty-hour workweek may be required.

Employee's Acknowledgement: I certify that I am able to perform, with or without the assistance of a reasonable accommodation, the essential duties of this position.

Employee's Name (Print): _____

Employee's Signature: _____ **Date:** _____

Supervisor's Acknowledgment: I certify this duty statement represents a current and accurate description of the essential functions of this position. I have discussed the duties of this position with and provided the above-named employee a copy of this duty statement.

Supervisor's Name (Print): Jason Harville

Supervisor's Signature: _____ **Date:** _____