

# **CSUB Math Field Day: Data Science Exploration Challenge**

## **1. Competition overview**

### **Audience**

High School students, grades 11 to 12.

### **Goal**

Teams of **three to four students** explore a real data set related to an annual theme, perform a meaningful analysis **using any tools (no restrictions)** they choose, and present their findings in a 10-minute talk followed by 2 minutes of questions.

### **Spirit of the event**

- Focus on curiosity, creativity, and communication, not on advanced statistical analysis.
- Encourage teamwork and inclusive participation from a wide range of schools.
- Emphasize statistical thinking, data visualization and clear storytelling

## **2. Annual theme and datasets**

Each year CSUB Math Department selects a theme (for example, Climate and Health, Sports Analytics, Education and Equity, Local Environment, etc).

### **Datasets for teams**

- CSUB faculty will prepare one cleaned dataset that matches the yearly theme.

### **Datasets will be:**

- Organized and ready to use, with basic cleaning already done
- Free of any personal identifying information
- Moderate in size, appropriate for high school work

All registered teams will use the same dataset for the competition. Within this dataset, each team is free to choose its own research question and the story it wants to tell. Different teams may explore different variables, relationships, or subgroups in the data.

### **Access to data and extra sources**

All teams receive access to the common dataset, and a short data dictionary (variable names, units, brief descriptions) on the dataset release date listed in the timeline below.

Teams may add publicly available data to complement the provided dataset if they wish. Any additional data must be clearly documented in the slides, including proper citations and links to the source. The primary focus of the project should remain on analyzing the CSUB provided dataset, with extra data used only to complement your findings, provide context, comparisons, or additional insight.

### 3. Timeline and milestones

**Google form for registration:** <https://forms.gle/hgAeXSsY37P2kYwu8>

- Registration deadline: **Friday, January 30.**
- Dataset and data dictionary emailed to registered teams: **Friday, February 6.**
- Final slide deck due: **Thursday, March 6**

Teams must email the final slide deck (PDF or PowerPoint) to CSUB:

**Email:** [ayatawara@csub.edu](mailto:ayatawara@csub.edu)

**Title: “CSUB Math Field Day: Data Science Exploration Challenge, Final presentation – High school name”**

### 4. Deliverable and presentation format

The only required deliverable is a slide deck for a 10-minute oral presentation during Math Field Day on **Friday, March 7**. Teams have 10 minutes for their presentation followed by 2 minutes for questions from judges and the audience. Check the end of the document for the **presentation structure and analysis guidelines**.

### 5. Judging and awards

Each team is evaluated by a panel of three judges. Each judge scores the team using the five category rubric (0 to 4 in each category), and the scores are combined to determine the final team score. Prizes are awarded for 1st, 2nd, and 3rd place overall

### 6. Rules and Code of Conduct

- Teams must consist of 3 to 4 students from the same school.
- All students must be enrolled in grades 11 or 12.
- All teams must use the CSUB-provided dataset as the primary data source.
- Any additional public data must be cited and clearly identified in slides.
- Work must be completed by students. Teachers may clarify instructions but may not direct analysis or slide creation.
- All analysis and presentation content must be the team’s own work.

- AI tools and software are allowed, but students must understand and be able to explain anything they present.

## **7. Presentation structure and analysis guidelines**

### **1. Title slide**

- Project title, theme, dataset chosen.
- School name and team member names.

### **2. Context and question**

- Brief background on the theme.
- Clear statement of the main question(s) or goal.
- Why this question matters.

### **3. Data description**

- What dataset was used.
- Key variables.
- Any filtering or simple cleaning they did.

### **4. Methods**

- Briefly describe the main analyses or models you used in clear, everyday language.

Typical tools might include:

- Comparing groups or categories
- Looking at associations (for example, correlation)
- Simple or multiple regression
- Time trend or time series patterns

You should start with ideas that are in the spirit of CP Statistics, AP Statistics, or dual enrollment Statistics courses. You may use methods beyond that level (for example, multiple regression or more advanced models) only if you can:

- Explain in your own words what the method does
- Justify why it is appropriate for your question
- Interpret its output correctly

- Coding and software are welcome for organizing data, running analyses, and making visualizations. They should not be used to add methods that the team cannot explain.

## 5. **Key statistical findings**

Summarize the most important numerical or statistical results that help answer your research question. Focus on:

- Clear interpretation in context rather than formulas or jargon
- Linking each result back to the question you posed
- Showing why your analysis is reasonable and trustworthy (validity of analysis)
- Highlighting any interesting, unexpected, or insightful patterns you discovered (insight)

Your audience should be able to understand your main takeaways if they are comfortable with AP Statistics ideas, even if you used a few more advanced tools.

## 6. **Visualizations**

- At least **two clear, well labeled graphs** that support their main points.
- Students explain what each graph shows and why it matters.

## 7. **Limitations and next steps**

- At least one limitation of their analysis or the data.
- Ideas for what they would do with more time or more data.

## 8. **Conclusion**

- Short recap of the main answer to their question.
- Practical or real-world implications if possible.

- Judges or moderator give a **2 minute warning**.
- Presentations must be between 9 and 11 minutes. Teams that go under 9 minutes or over 11 minutes receive a small deduction in the Delivery, teamwork and Q&A category.