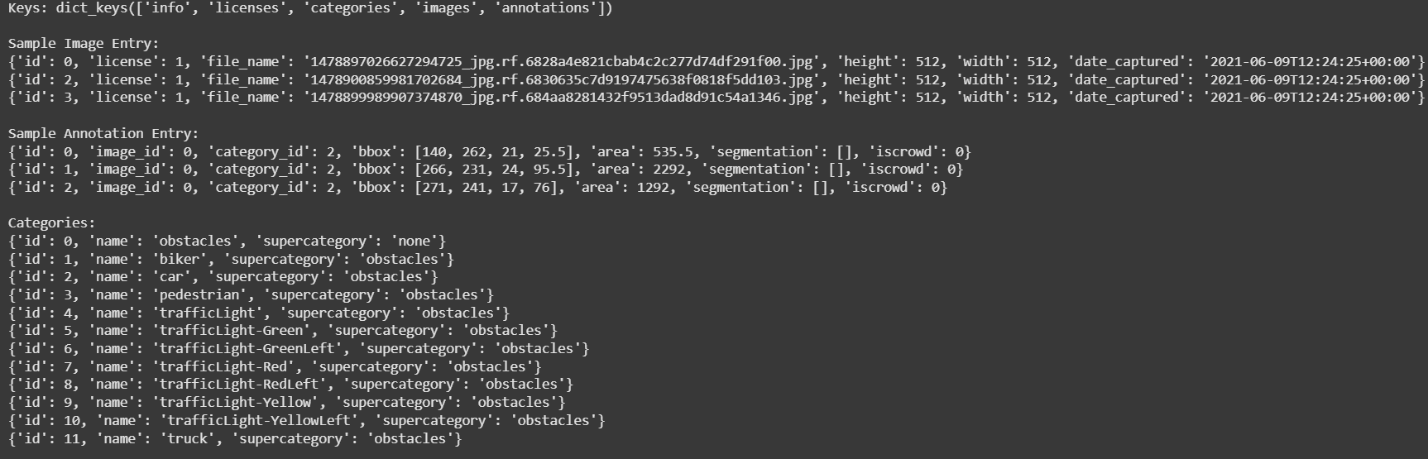
# Dataset Exploration Report

**1. Introduction**

* **Dataset Source:** COCO (Common Objects in Context)
* **Purpose:** Object detection in autonomous driving scenarios, aiming to identify and classify objects like pedestrians, vehicles, traffic signs, and obstacles in real-time.

**2. Dataset Composition**

* **Number of Images:** 26300 (deleted unannotated images).
* **Number of Objects (Bounding Boxes):** 11.
* **Classes:**



* **Bounding Box Format:** COCO format [x\_min, y\_min, width, height].

**3. Class Distribution Analysis (📊)**

* **Histogram of Object Categories:** A visualization showing the distribution of object categories to check for class imbalances. **A graph with blue squares

  AI-generated content may be incorrect.**

**4. Image Quality & Metadata Analysis**

* **Resolution Distribution:**

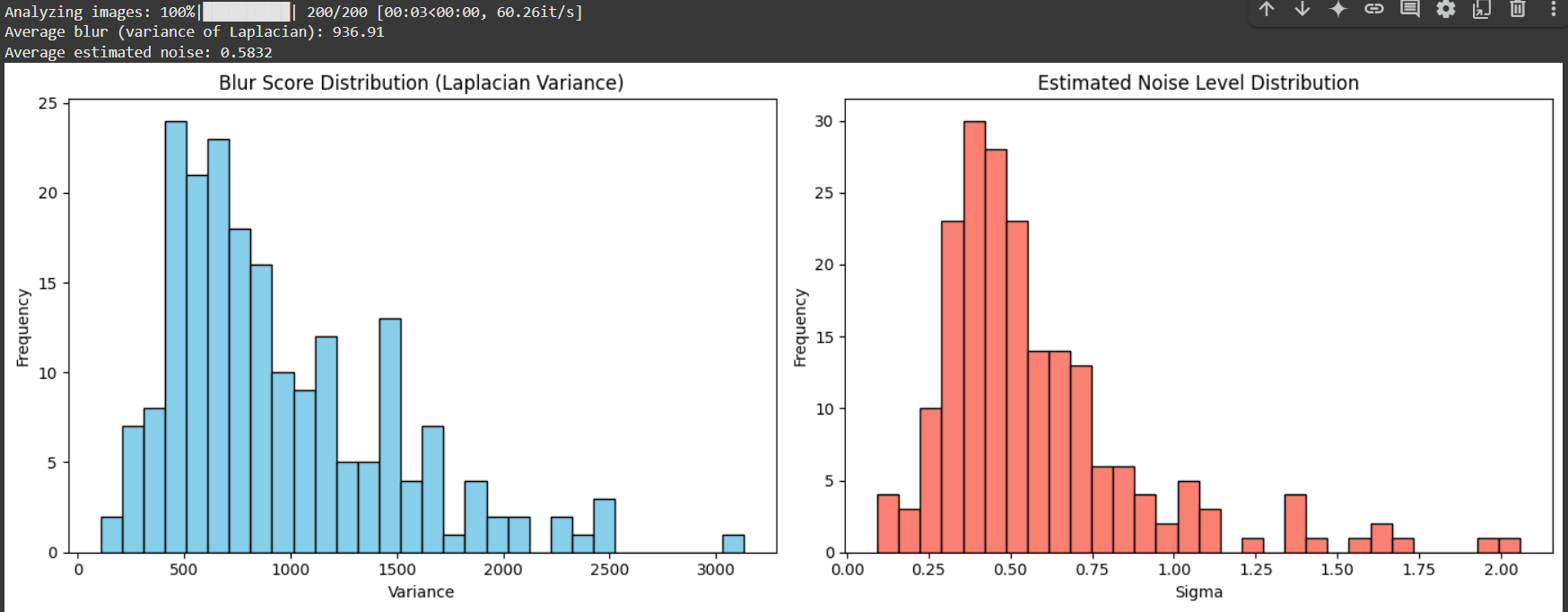
**Image Resolution Stats:**

* **Min size: 512x512**
* **Max size: 512x512**
* **Avg size: 512x512**

A graph with a red line

AI-generated content may be incorrect.

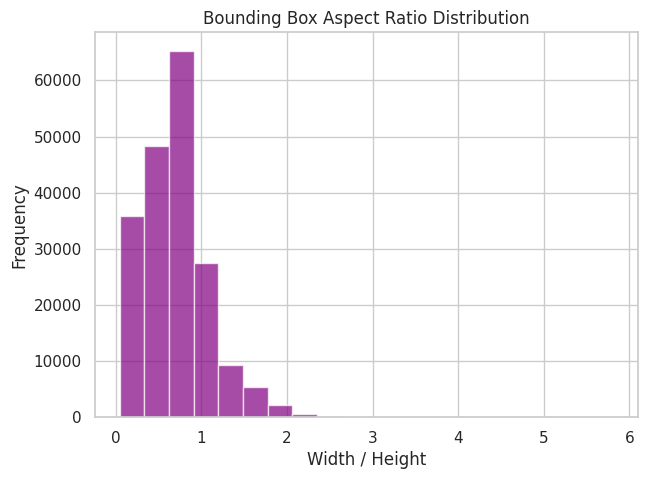
* **Blur & Noise Analysis:** Identification of images with low quality due to motion blur, poor lighting, or noise.



**5. Bounding Box Validity Check**

* **Total Bounding Boxes:** 194539
* **Invalid Bounding Boxes:** Zero (number and percentage of bounding boxes exceeding image dimensions or having negative coordinates).
* **Bounding Box Size Distribution:** Statistics on bounding box widths and heights. A graph of a number of different sizes

  AI-generated content may be incorrect.



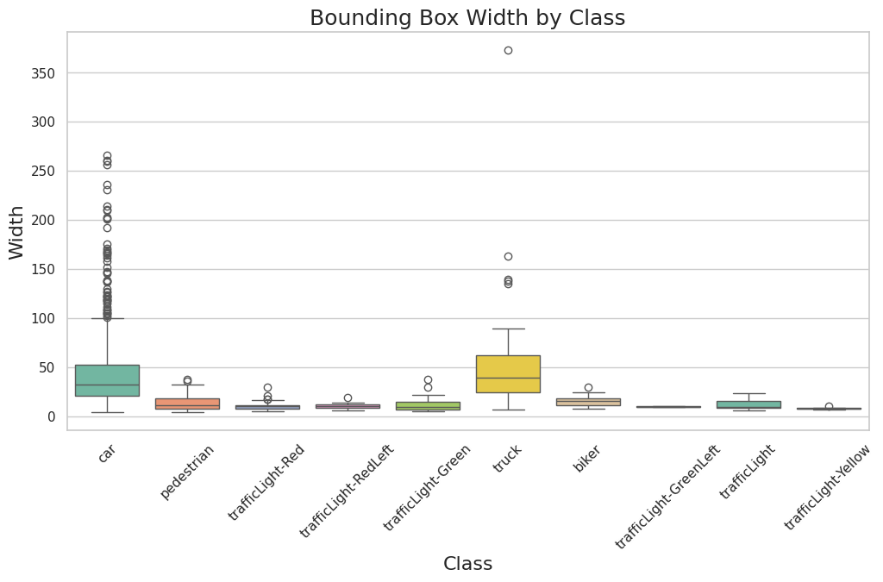
**6. Spatial Distribution Analysis**

* **Spatial Distribution Analysis**

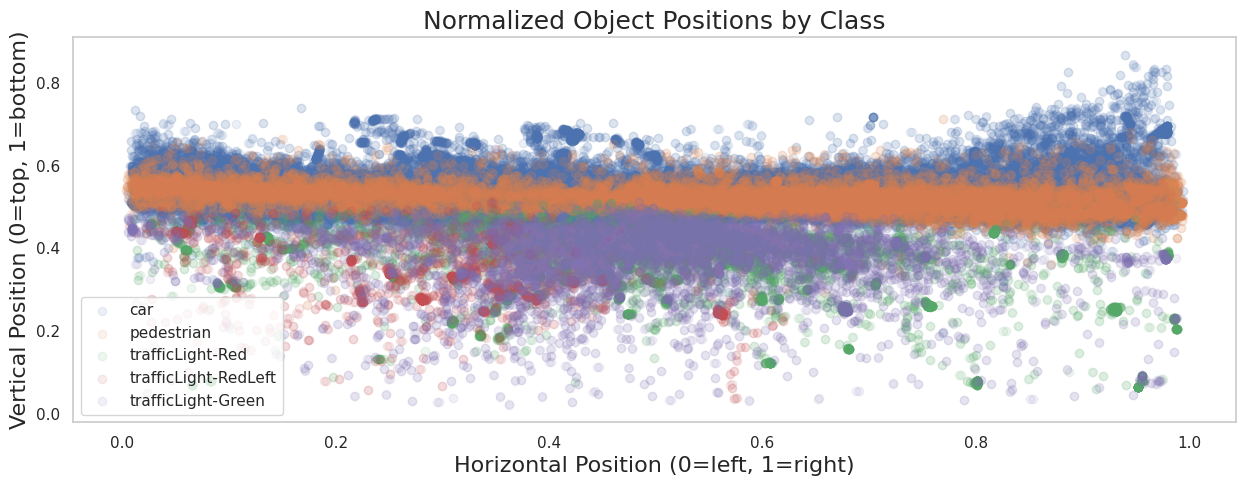
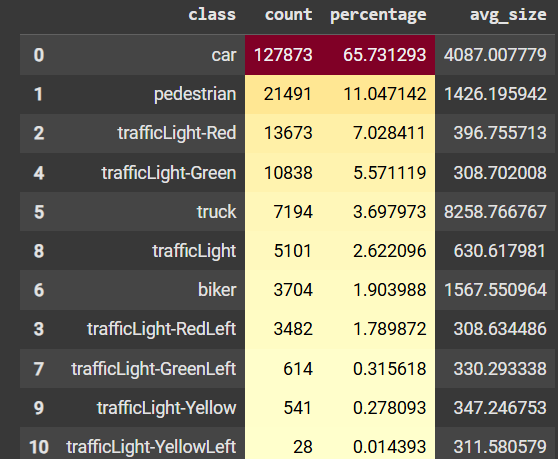
**A map of heatmap

AI-generated content may be incorrect.**

* **Bounding Box Width by Class**



* 1. **Bias Detection & Visualization:**

****