Hi there, this is Group 12. Our team members include Yike, Kehan, Calvin, and myself, Tianao. Let's start with the data source, corresponding to Task 1. After consulting relevant literatures, we dentified this four columns that determine the diversity of the TCR gene and identified antigen.gene as the target column for prediction. We also retained columns related to MHC, which help us understand TCR specificity. The vdjdb.score column represents the quality score of the data records. It can serve as a basis for data filtering or as a reference for weighting. Next, we visualized these columns—not just basic statistics but also cross-analysis of the column score and the column antigen with other columns. For example, this chart is 'Distribution of Gene Types for Top 10 Antigen Epitopes'.