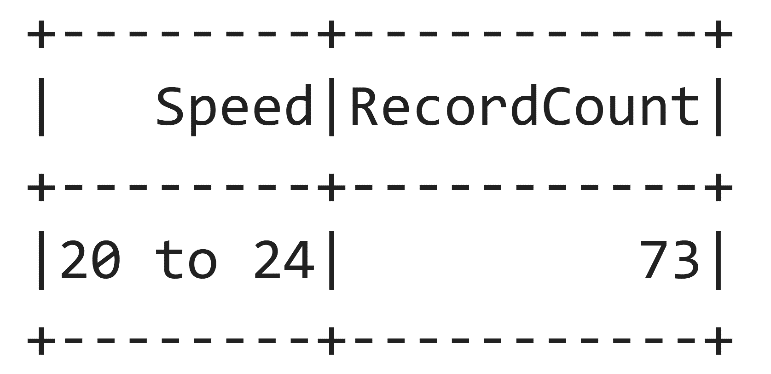
**Determine what vehicle speed was present the most during crashes.**

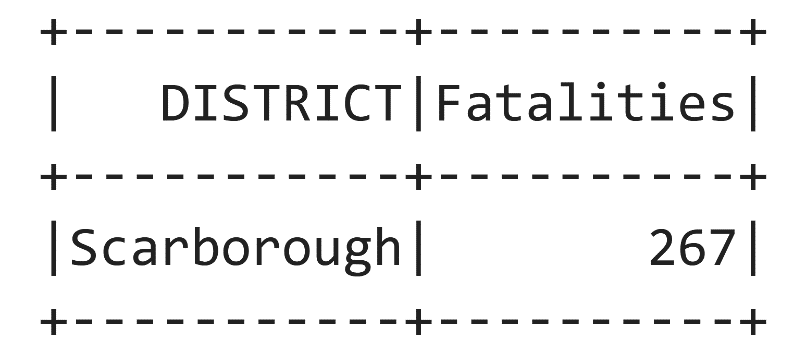
spark.sql("SELECT INVAGE AS Speed, COUNT(\*) AS RecordCount FROM fatalities GROUP BY INVAGE ORDER BY RecordCount DESC LIMIT 1").show()



This goal aims to analyze and identify the predominant vehicle speed associated with crashes. By examining crash data, the project will provide insights into the most common or prevalent speed at the time of accidents, offering valuable information for understanding patterns and potential contributing factors to improve road safety measures.

**Determine what district had the most fatal car crashes.**

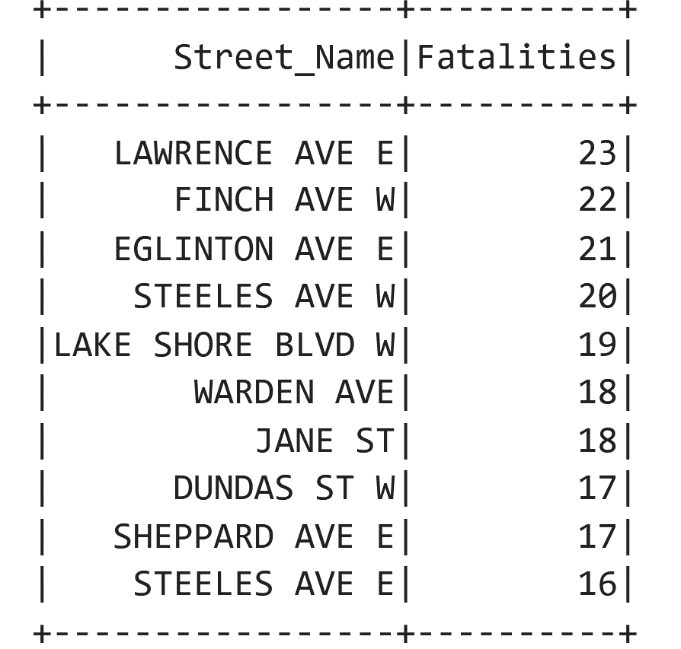
spark.sql("SELECT DISTRICT AS DISTRICT, COUNT(\*) AS Fatalities FROM fatalities GROUP BY DISTRICT ORDER BY Fatalities DESC LIMIT 1").show()



This goal was implemented to determine which district experienced the highest number of fatal car crashes. By analyzing relevant data, the project aims to identify geographic patterns and areas with elevated risks. This information can be crucial for directing targeted interventions, implementing safety measures, and formulating policies to reduce the incidence of fatal car accidents in specific districts.

**Determine what street had the most fatal crashes.**

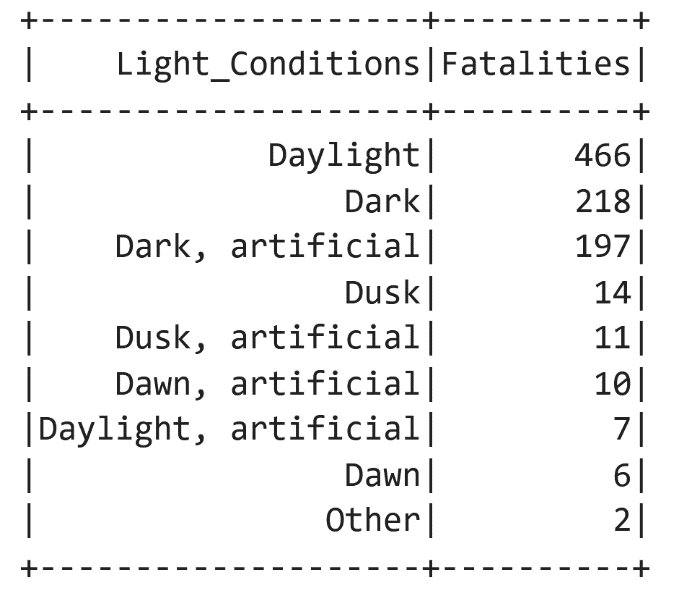
spark.sql("SELECT STREET1 as Street\_Name, count(\*) AS Fatalities FROM fatalities GROUP BY Street\_Name ORDER BY Fatalities DESC LIMIT 10").show()



This objective aims to identify the street that recorded the highest number of fatal crashes. By analyzing relevant data, the project will pinpoint specific locations with elevated risks, allowing for a focused approach in implementing safety measures, traffic management strategies, and targeted interventions to enhance road safety on the identified street.

**Determine how many fatal crashes occurred during the day as opposed to during night.**

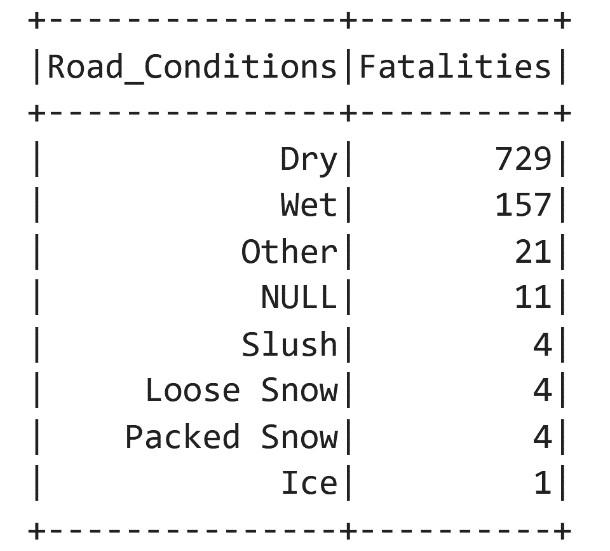
spark.sql("SELECT LIGHT as Light\_Conditions, count(\*) AS Fatalities FROM fatalities GROUP BY Light\_Conditions ORDER BY Fatalities DESC").show()



This goal involves analyzing data to quantify the number of fatal crashes that occurred during the day and night, as well as different artificial lighting conditions. By distinguishing between daytime and nighttime incidents, the project aims to provide insights into temporal patterns of fatal crashes. This information can be valuable for understanding potential factors contributing to accidents during specific periods, guiding efforts to improve safety measures, and enhancing overall road safety strategies.

**Determine how many fatal crashes occurred while roads were wet as opposed to dry.**

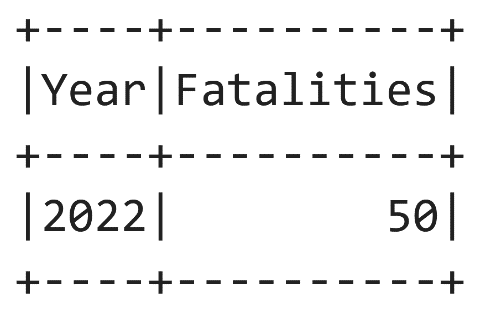
spark.sql("SELECT RDSFCOND AS Road\_Conditions, count(\*) AS Fatalities FROM fatalities GROUP BY Road\_Conditions ORDER BY Fatalities DESC").show()



This objective involves analyzing data to quantify the number of fatal crashes that occurred during wet road conditions compared to dry conditions. By distinguishing between these weather conditions, the project aims to provide insights into the impact of road surface conditions on fatal crashes. This information can be crucial for understanding the relationship between weather-related factors and road safety, guiding efforts to implement weather-specific safety measures and enhance overall road safety strategies.

**Determine what year had the most fatal car crashes.**

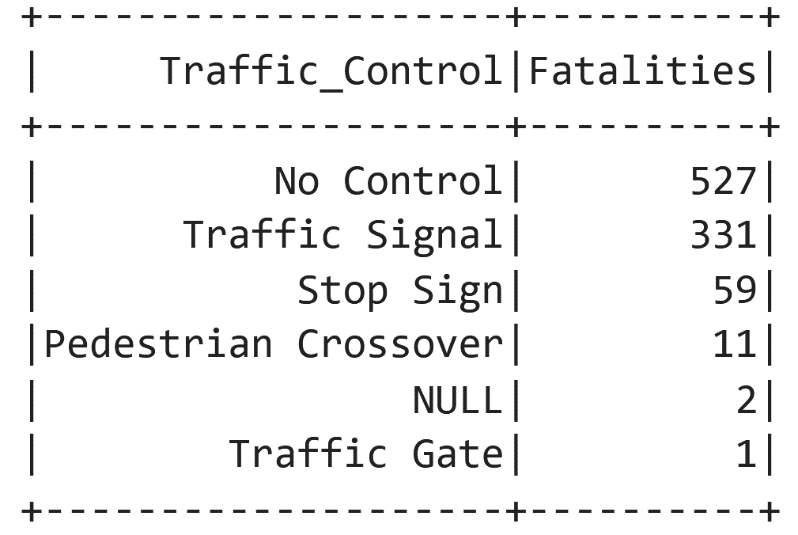
spark.sql("SELECT YEAR as Year, count(\*) AS Fatalities FROM fatalities GROUP BY Year ORDER BY Year DESC LIMIT 1").show()



This goal involves analyzing historical data to identify the year with the highest number of fatal car crashes. By examining trends over time, the project aims to pinpoint specific years that experienced elevated risks, providing valuable insights into potential contributing factors. This information can be essential for understanding temporal patterns in road safety and may help guide efforts to address factors associated with the identified year of peak fatal crashes.

**Determine what traffic control had the most fatal collisions.**

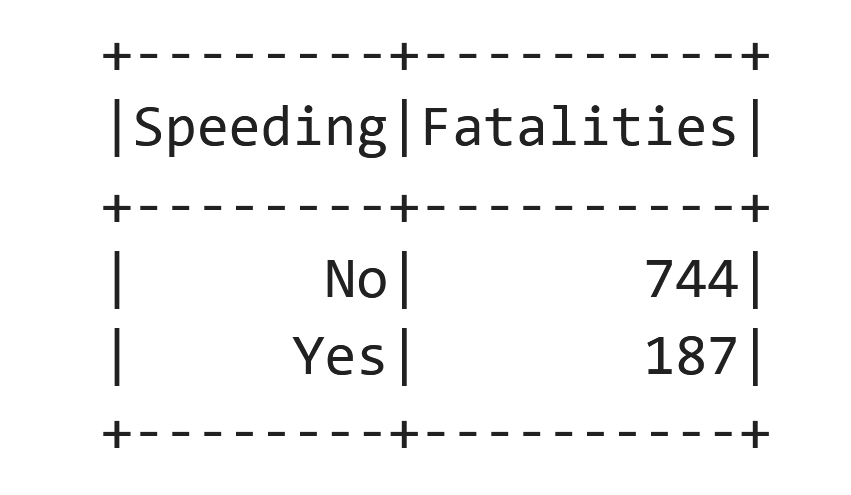
spark.sql("SELECT TRAFFCTL AS Traffic\_Control, count(\*) AS Fatalities FROM fatalities GROUP BY Traffic\_Control ORDER BY Fatalities DESC").show()



This objective involves analyzing data to identify the type of traffic control (e.g., stop signs, traffic signals, yield signs) associated with the highest number of fatal collisions. By examining patterns related to different traffic control measures, the project aims to provide insights into the effectiveness or vulnerabilities of specific control systems. This information can be valuable for enhancing traffic management strategies, improving signage, and implementing targeted interventions to reduce fatal collisions at intersections with the identified traffic control.

**Determine how many fatal collisions occurred while driver exceeding the speed limit.**

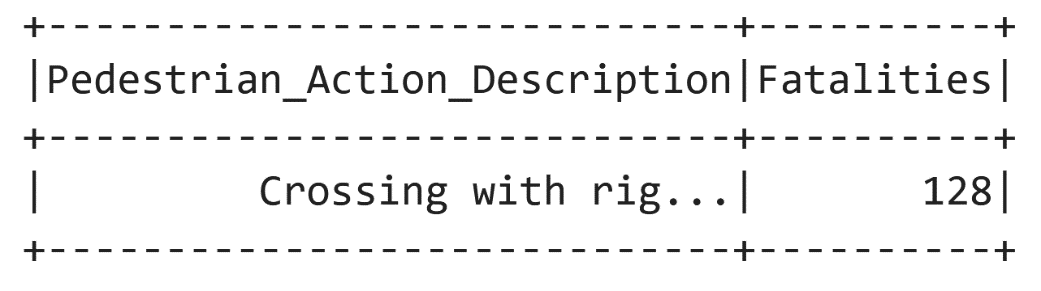
spark.sql("SELECT SPEEDING AS Speeding, count(\*) AS Fatalities FROM fatalities GROUP BY Speeding ORDER BY Fatalities DESC").show()



This goal involves analyzing data to quantify the number of fatal collisions that occurred when drivers were exceeding the speed limit. By examining the correlation between speed violations and fatal accidents, the project aims to provide insights into the impact of speeding on road safety. This information can be crucial for understanding the relationship between speed-related factors and fatal collisions, guiding efforts to enforce speed limits, implement targeted interventions, and enhance overall strategies for reducing speeding-related fatalities.

**Determine the count of fatal collisions that occurred while pedestrian crossing the right of way.**

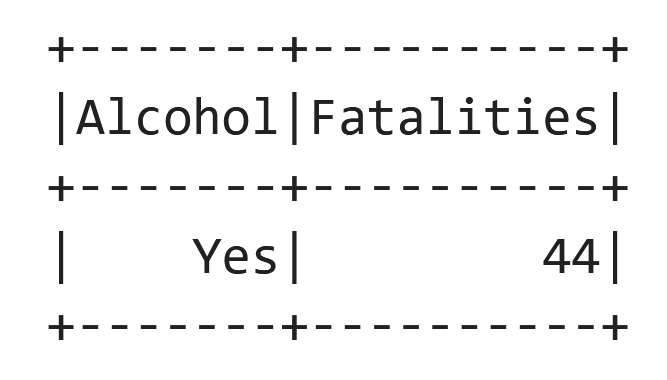
spark.sql("SELECT PEDACT AS Pedestrian\_Action\_Description, count(\*) AS Fatalities FROM fatalities WHERE PEDACT='Crossing with right of way' GROUP BY Pedestrian\_Action\_Description").show()



This objective involves analyzing data to determine the count of fatal collisions that occurred while pedestrians were crossing the right of way. By examining incidents where pedestrians were following proper crossing procedures, the project aims to provide insights into pedestrian safety and potential risks associated with crossing designated areas. This information can be valuable for understanding the dynamics of pedestrian-involved fatal collisions, guiding efforts to enhance pedestrian safety measures, and informing public awareness campaigns to reduce accidents in crosswalks and designated pedestrian areas.

**Determine the drivers' state of condition for the most fatal collisions.**

spark.sql("SELECT ALCOHOL AS Alcohol, count(\*) AS Fatalities FROM fatalities WHERE ALCOHOL='Yes' GROUP BY Alcohol").show()



This goal involves analyzing data to determine the drivers' state of condition associated with the most fatal collisions. By examining driver impairment, the project aims to provide insights into the primary contributing factors to fatal collisions. This information can be crucial for understanding the role of driver behavior and conditions in road safety, guiding efforts to address specific issues, and informing strategies for promoting safer driving practices to reduce the number of fatal collisions.