Een Data Scientist heeft verschillende taken. Zo probeert hij relaties tussen verschillende variabelen te vinden. Dit doet hij bijvoorbeeld, zoals in opdracht 10, met een bivariate analysis op 2 numerieke attributen. Hierbij vind hij de correlatie tussen 2 variabelen en ziet hij de relaties ertussen. Daarna kan de data scientist, zoals in opdracht 18, een regression tree maken om voorspellingen te doen op basis van deze relaties. Hierbij probeert de Data Scientist een model te maken zodat bij het vinden van een nieuwe diamant, je bijvoorbeeld de prijs van de diamant kan voorspellen op basis van een van te voren besloten attribuut, in dit geval karaat.