Abstract: We model a market for data in which different datasets can be combined to create value in flexible ways. This gives rise to a graph representation of the market that captures rich patterns of substitutability and complementarity. We analyze how these patterns determine competitive outcomes, incentives to merge and incentives to acquire data. Equilibria can be inefficient but bundling restores allocative efficiency. The question ``who competes with whom'' is non-trivial. We argue that effective competition policy requires an understanding of the complement-substitute graph that our model develops.