1. Τα δείγματα του μικρού dataset (S06) αφορούν σε ανθρώπους που είναι γνωστό για το εάν έχουν (ή όχι) χρεωθεί πιστωτική κάρτα.

2. Υπάρχουν καταχωρημένα 690 δείγματα το καθένα με 15 χαρακτηριστικά (συν ενός για την κατηγορία του instance) που είναι κατηγορικά και αριθμητικά (real, integer).

3. Δεν υπάρχουν επικεφαλίδες, ούτε αρίθμηση γραμμών.

4. Οι κλάσεις έχουν ως ετικέτα είτε το πρόσημο (+) που δείχνει ότι ο εκάστοτε ερωτηθείς έχει χρεωθεί πιστωτική είτε το πρόσημο (−) ως αντίθετη ένδειξη. Βρίσκονται στην 16η στήλη (τελευταία κολόνα) .

5.

6. Υπάρχουν απουσιάζουσες τιμές σε 37 εκ των δειγμάτων, ήτοι 5% επί του συνόλου.

7. Υπάρχουν δύο κλάσεις, η μεν κλάση με label (+) και ποσοστό δειγμάτων 44.5% επί του συνόλου, η δε κλάση με label (−) και ποσοστό δειγμάτων 55.5% επί του συνόλου. Παρατηρούμε ότι η κλάση (−) είναι ~1.219 φορές πιο συχνή από την κλάση (+), οπότε σύμφωνα με την εκφώνηση το δείγμα θεωρείται ισορροπημένο.