

Γιατί R

- Ισχυρή γλώσσα προγραμματισμού.
- Εύχρηστη, δυναμική γλώσσα και εξαιρετικά ευέλικτη.
- Ότι καλύτερο διατίθεται σήμερα για στατιστική ανάλυση· μάλιστα η χρήση της R στη στατιστική αυξάνεται σημαντικά τα τελευταία χρόνια.
- Πολλές δημοσιεύσεις που προέρχονται από το χώρο της στατιστικής επιστήμης, παρέχουν την ανάλυσή τους σε μορφή R, έτσι ότι καινούργιο δημοσιεύεται καθίσταται άμεσα προσβάσιμο όχι μόνο στους θεωρητικούς της στατιστικής, αλλά και όσους χρησιμοποιούν στατιστικές μεθόδους στην έρευνά τους.

Γιατί R

- Διαθέτει δυνατότητες προγραμματισμού, που παρέχουν εξαιρετική ευελιξία στον ερευνητή, πέρα από τις βασικές ανάγκες της στατιστικής ανάλυσης. Μαθαίνοντας ο ερευνητής την R εισάγεται σε ένα καινούργιο κόσμο δυνατοτήτων, και καθίσταται ικανός να αξιοποιεί στο μέγιστο βαθμό τις υπολογιστές δυνατότητες που του παρέχουν οι σύγχρονοι ηλεκτρονικοί υπολογιστές.
- Η R διαθέτει ένα επιπλέον πλεονέκτημα σε σχέση με τα υπόλοιπα στατιστικά πακέτα, είναι ανοικτού κώδικα, αυτό σημαίνει ότι κανείς μπορεί να γνωρίζει τι συμβαίνει όταν διεξάγει μία ανάλυση εξετάζοντάς τον κώδικά της.
- Πολύ καλό γραφικό περιβάλλον.
- Επίσης, ενώ τα περισσότερα πακέτα στατιστικής ανάλυσης κοστίζουν πάρα πολλά λεφτά για να τα αγοράσει κανείς, η R είναι εντελώς δωρεάν.

Γιατί R

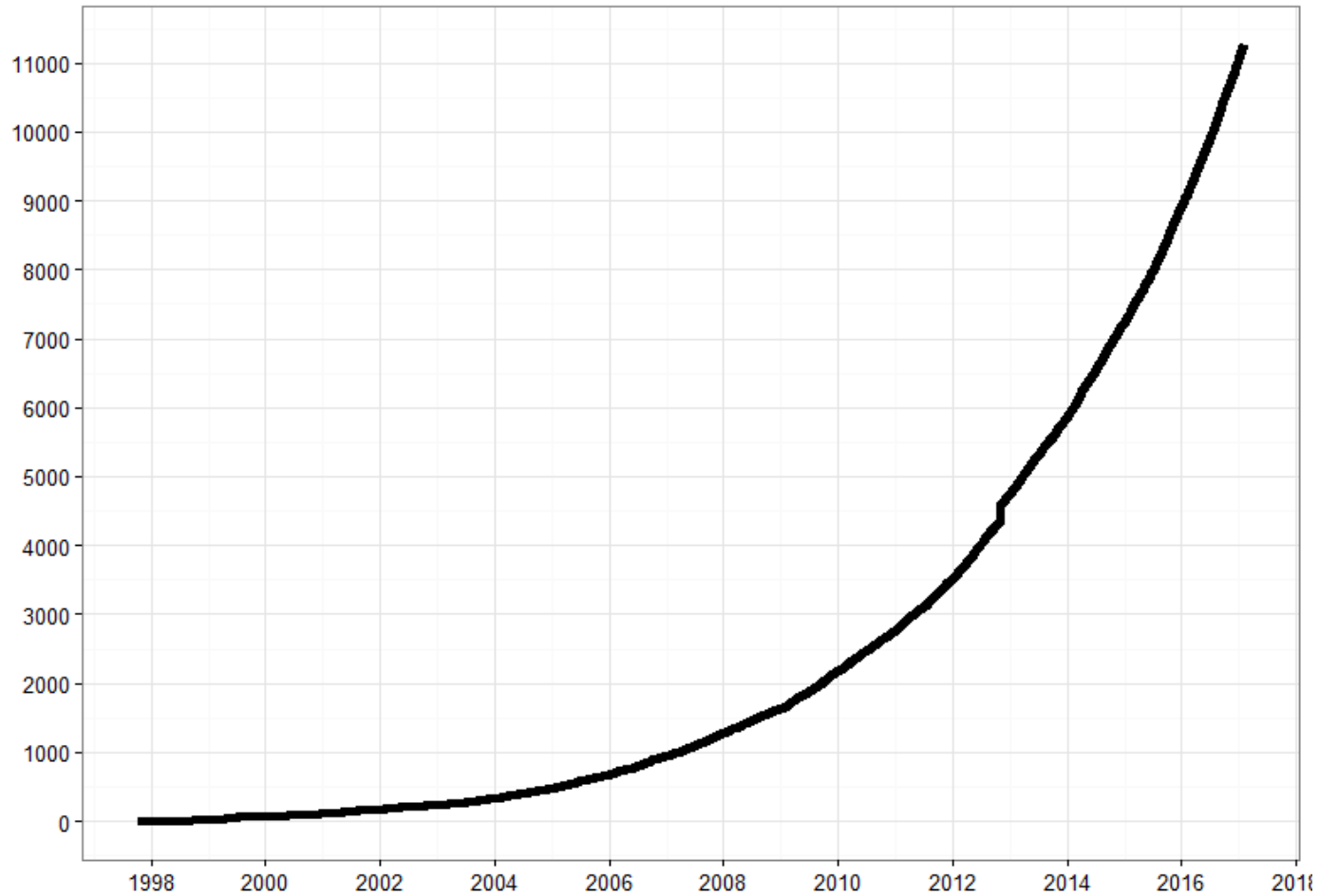
- Η R έρχεται με μια σειρά πακέτων (packages) που προσφέρουν βασικές στατιστικές λειτουργίες. Τα πακέτα αυτά τα καλεί ο χρήστης από τη γραμμή εντολών του προγράμματος (το R δεν διαθέτει κάποιο ξεχωριστό γραφικό περιβάλλον (GUI) και μενού εντολών για τις στατιστικές αναλύσεις) και στη συνέχεια έχει πρόσβαση στις ρουτίνες του κάθε πακέτου. Εκτός όμως από τα βασικά πακέτα που διανέμονται με το πρόγραμμα, η κοινότητα του R έχει συνεισφέρει συνολικά πάνω από 10000 πακέτα που επεκτείνουν τις ήδη ευρύτατες δυνατότητες του προγράμματος με εφαρμογές στις κοινωνικές επιστήμες, στην οικονομική ανάλυση, στη φασματοσκοπία, στη ανάλυση γονιδιωμάτων και αλλού. Οι εφαρμογές είναι πραγματικά απεριόριστες.

Companies Using R

- Google uses R to calculate the ROI on advertising campaigns.
- Ford uses R to improve the design of its vehicles.
- Twitter uses R to monitor user experience.
- The US National Weather Service uses R to predict severe flooding.
- The Rockefeller Institute of Government uses R to develop models for simulating the finances of public pension funds.
- The Human Rights Data Analysis Group uses R to quantify the impact of war.
- R is used frequently by The New York Times to create infographics and interactive data journalism applications.

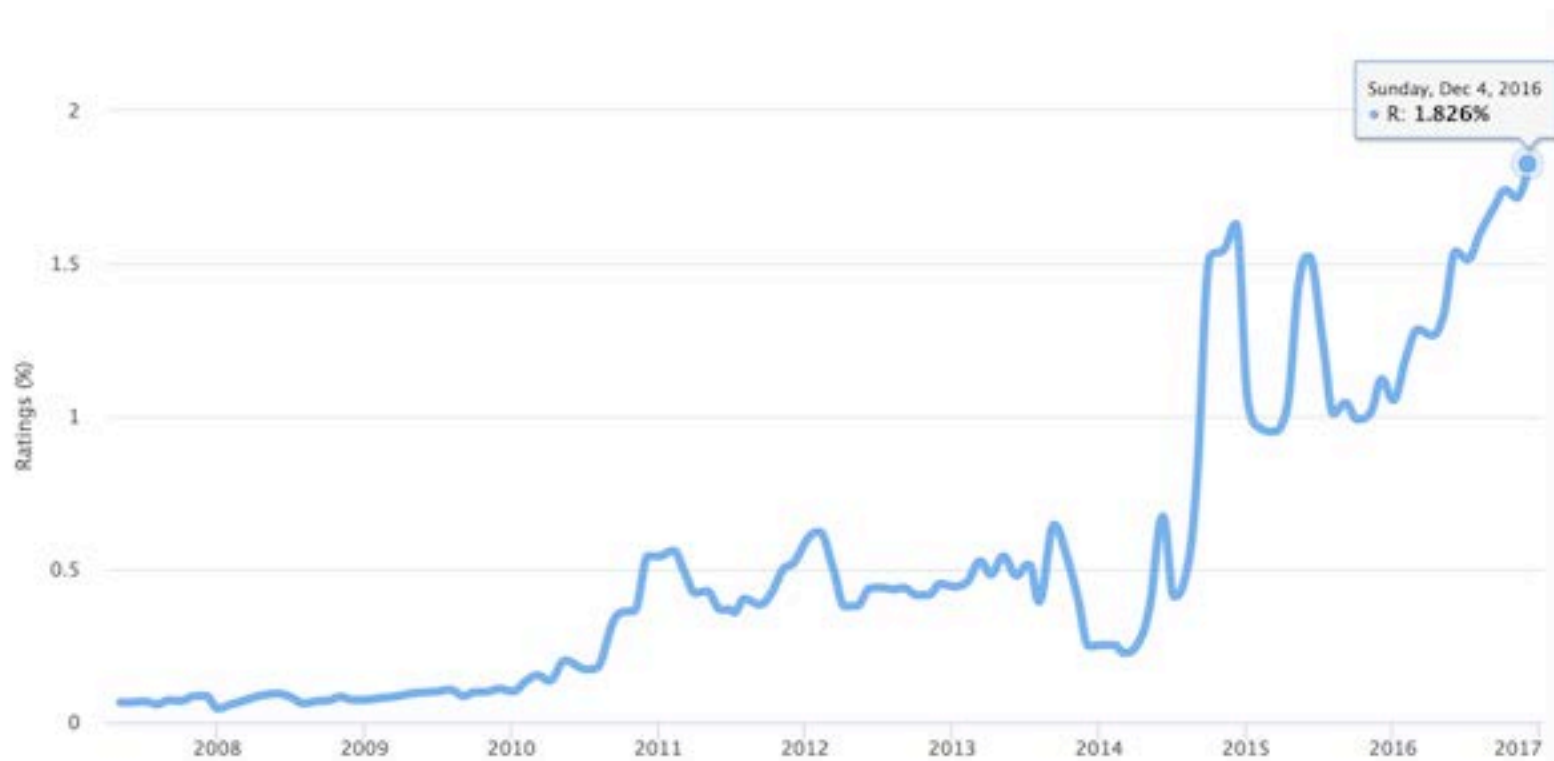
Packages

Number of R packages ever published on CRAN



Popularity

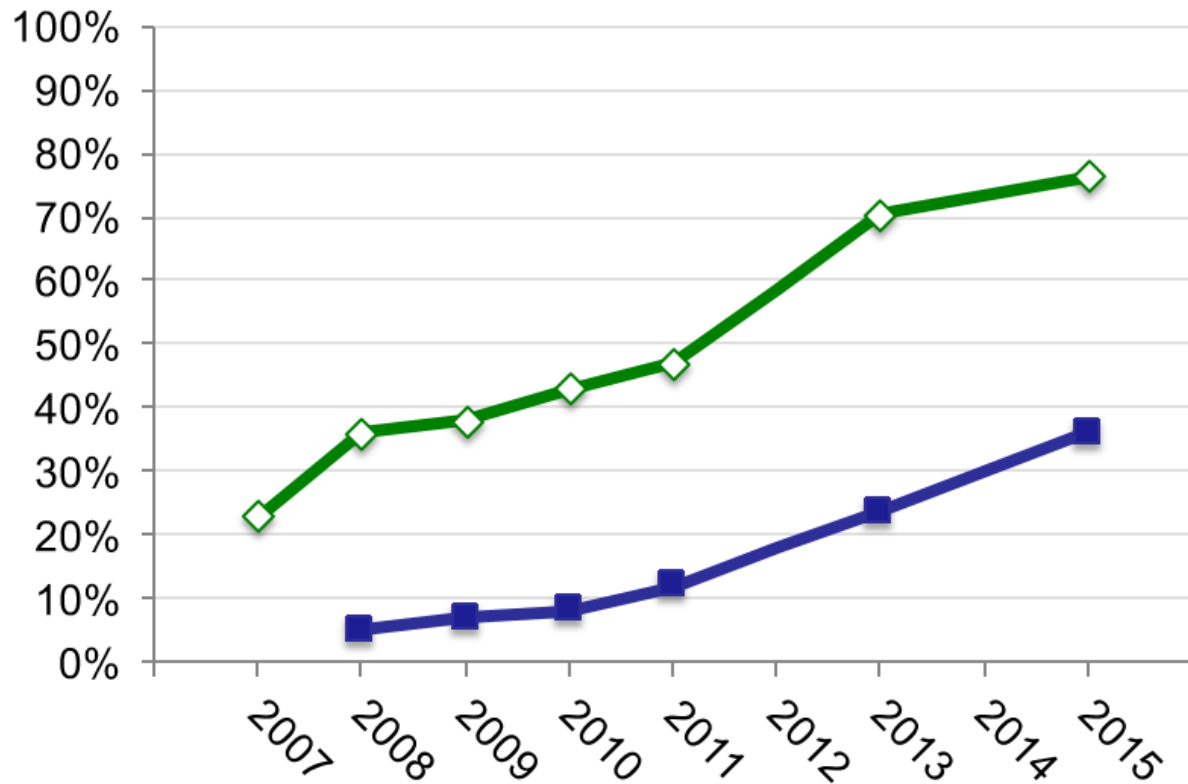
R'S POPULARITY HAS INCREASED DRAMATICALLY



source: www.tiobe.com

Popularity

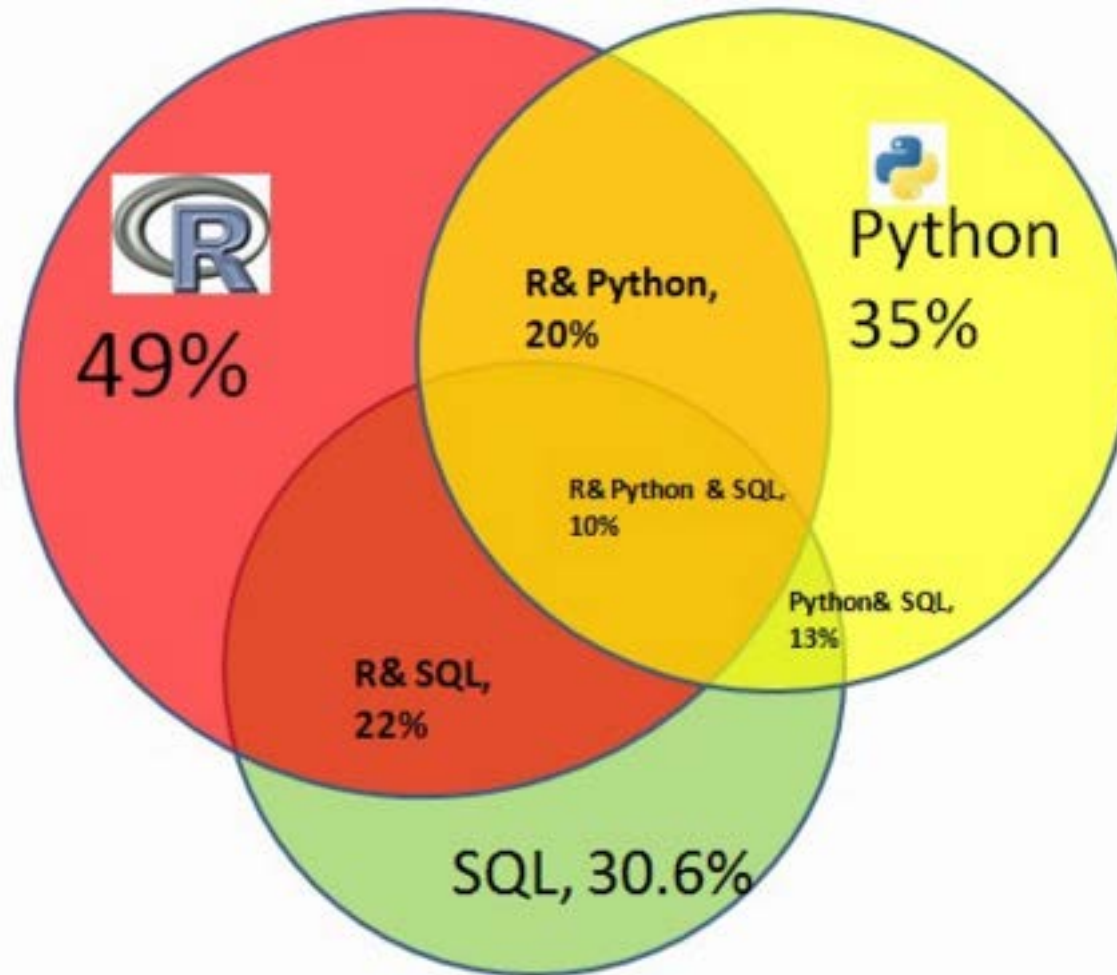
R Usage



76% of analytic professionals report using R

36% select R as their primary tool

R & Python in Data Science



R & Python in Data Science

- Python or R both is not rocket science. If you are a beginner in data science field with programming skill, the learning curve will be same for you for both Python and R. In contrast, if you are from a statistical background with no programming knowledge, R is the best choice for you. Definitely choice depends on goal as well. If your goal is not to become a data scientist and just to gain knowledge of data analysis for programming purpose Python is the best for you.
- From problem perspective, both R and Python stand on the same height for data analysis. In fact, performing data analysis in either language is almost similar. However, as a flexible and multipurpose data analysis tool Python wins the game.
- **A data scientist must be a blend of coder and statistician.** If you are already a programmer or your data science learning goal is for application development purpose; then Python is the right learning path for you. **But for a core data scientist definitely, there is no better choice than R.**
- **Keeping all the above data points in mind, finally, Python or R is your choice!**