

Wprowadzenie do eR(k)a

Bartosz Sękiewicz

b.sekiewicz@erkakrakow.pl



Learn more at
revolutionanalytics.com/what-is-r



This video is free to remix and share
What is R? by Revolution Analytics CC BY-SA 3.0

REVOLUTION
ANALYTICS

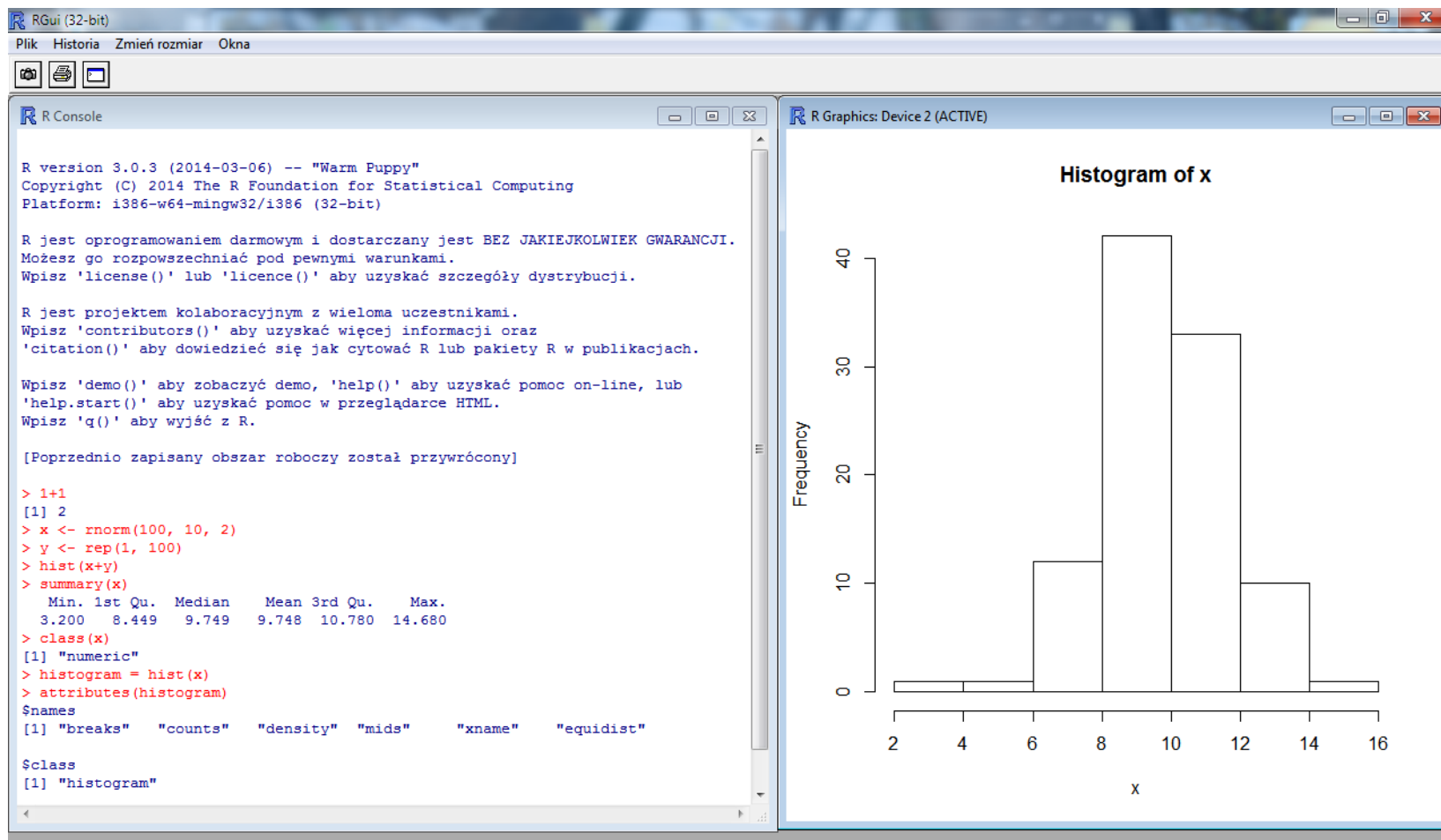
R w pigułce

- Co to jest R?
 - Pakiet statystyczny
 - Język programowania
 - Platforma programistyczna z interpreterem tego języka
 - Nazwa projektu, w ramach którego rozwijany jest język oraz środowisko
 - R to społeczność
- 1993 - Robert Gentleman oraz Ross Ihaka (R&R)
- 1997 - R Development Core Team (około 20 osób)
- 2002 - The R Foundation for Statistical Computing (setki uczestników)

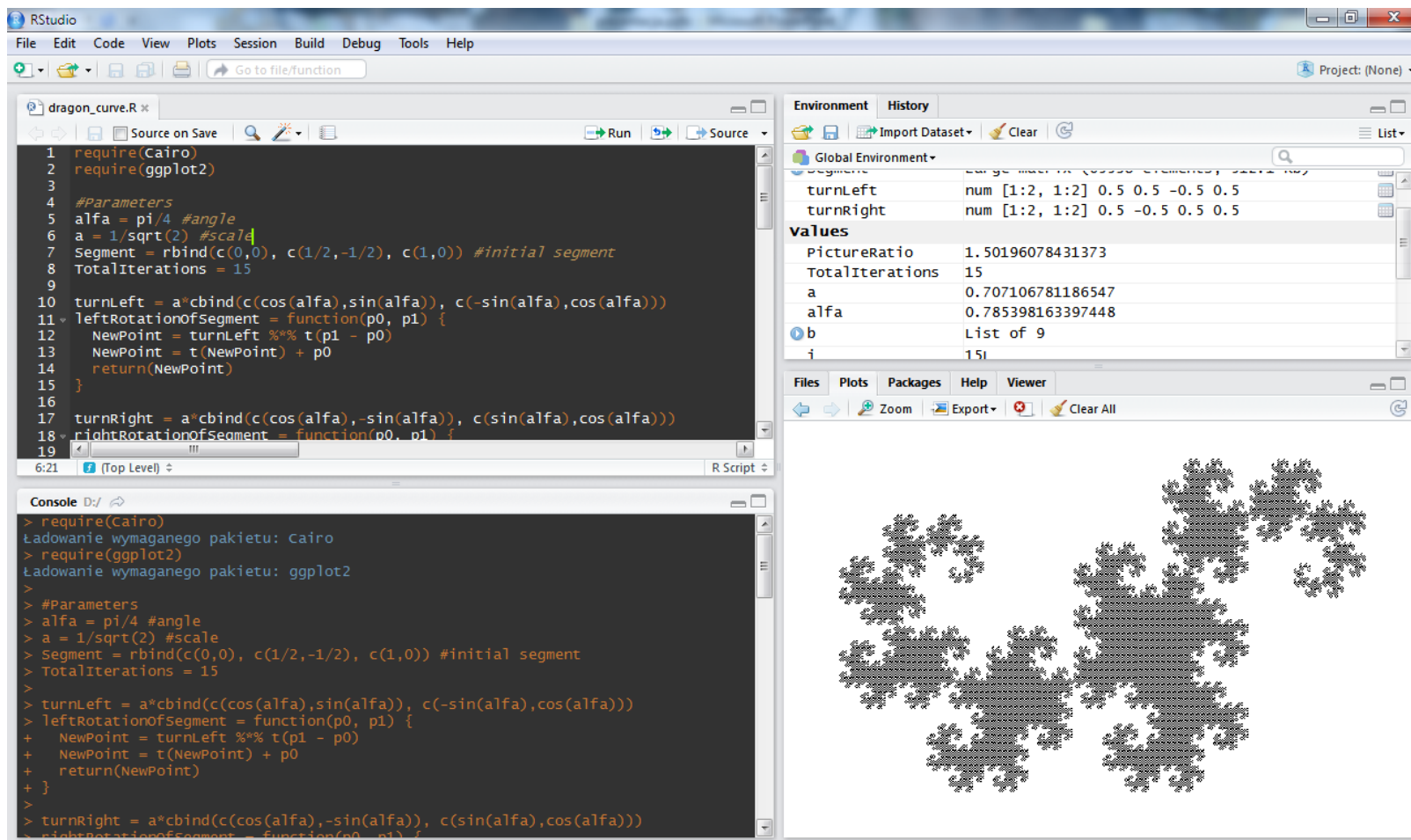
R to bardzo tanie rozwiązanie

- GNU R było, jest i zawsze będzie darmowe, większość pakietów także jest za darmo
- Narzędzia wspierające pisanie w R są za darmo (RStudio, które jest potężnym i wydajnym interfejsem użytkownika R, wtyczka NppToR dla Notepad++ itd.)
- System R działa praktycznie na wszystkich platformach

R GUI

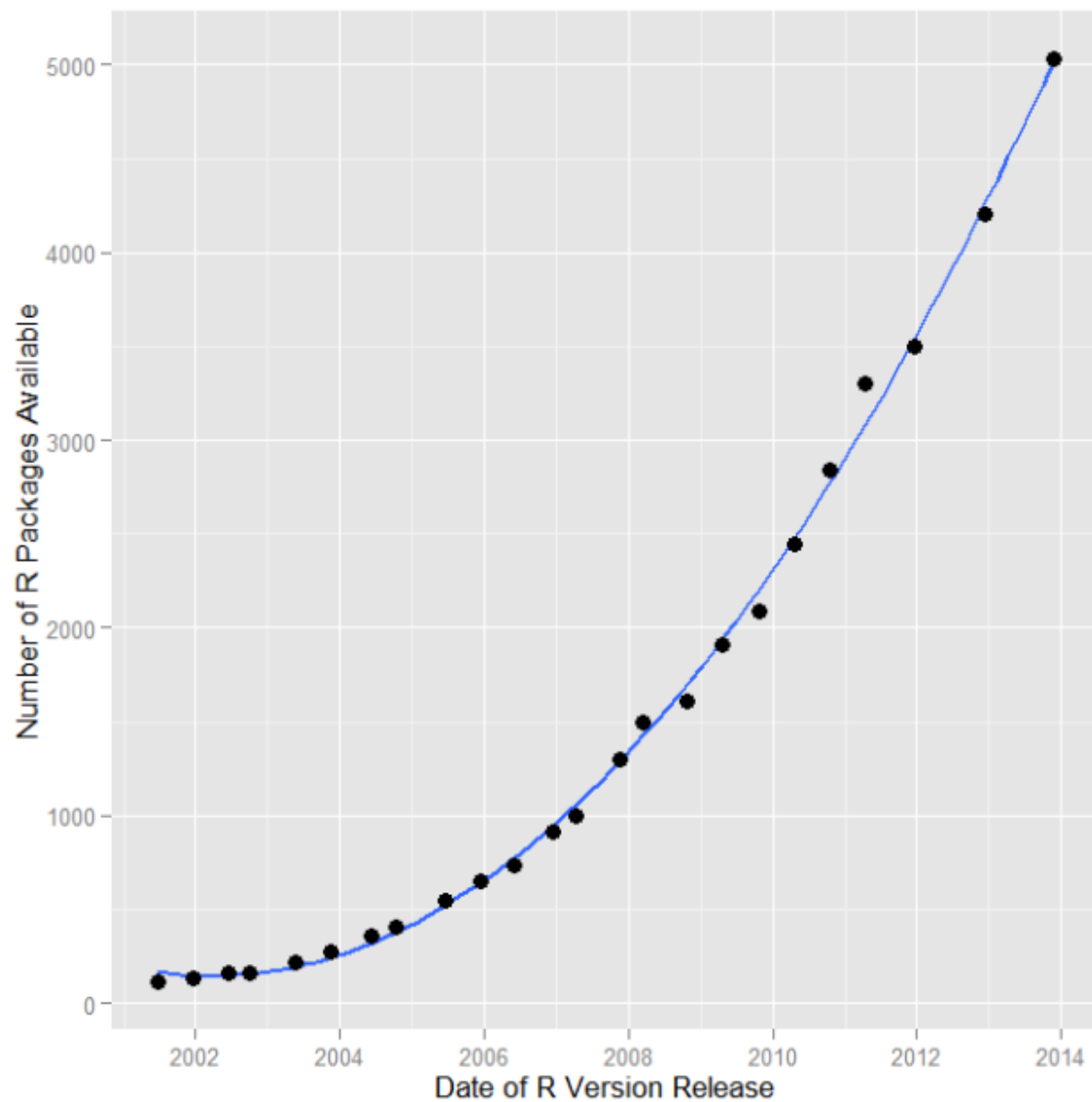


RStudio



R to pakiety

- Siła R tkwi w tysiącach pakietów tworzonych przez społeczność R
 - [CRAN](#) (The Comprehensive R Archive Network)
 - [Bioconductor](#)
 - Własne repozytoria (np. GitHub)
- Ponad 2 miliony użytkowników powoduje, że zarówno kod źródłowy, jak i popularne pakiety są ciągle sprawdzane i poprawiane



Bayesian	Bayesian Inference
ChemPhys	Chemometrics and Computational Physics
ClinicalTrials	Clinical Trial Design, Monitoring, and Analysis
Cluster	Cluster Analysis & Finite Mixture Models
DifferentialEquations	Differential Equations
Distributions	Probability Distributions
Econometrics	Computational Econometrics
Environmetrics	Analysis of Ecological and Environmental Data
ExperimentalDesign	Design of Experiments (DoE) & Analysis of Experimental Data
Finance	Empirical Finance
Genetics	Statistical Genetics
Graphics	Graphic Displays & Dynamic Graphics & Graphic Devices & Visualization
HighPerformanceComputing	High-Performance and Parallel Computing with R
MachineLearning	Machine Learning & Statistical Learning
MedicalImaging	Medical Image Analysis
MetaAnalysis	Meta-Analysis
Multivariate	Multivariate Statistics
NaturalLanguageProcessing	Natural Language Processing
NumericalMathematics	Numerical Mathematics
OfficialStatistics	Official Statistics & Survey Methodology
Optimization	Optimization and Mathematical Programming
Pharmacokinetics	Analysis of Pharmacokinetic Data
Phylogenetics	Phylogenetics, Especially Comparative Methods
Psychometrics	Psychometric Models and Methods
ReproducibleResearch	Reproducible Research
Robust	Robust Statistical Methods
SocialSciences	Statistics for the Social Sciences
Spatial	Analysis of Spatial Data
SpatioTemporal	Handling and Analyzing Spatio-Temporal Data
Survival	Survival Analysis
TimeSeries	Time Series Analysis
WebTechnologies	Web Technologies and Services
gR	gRaphical Models in R

<http://cran.r-project.org/web/views/>

Społeczność R

- Duża liczba stron w języku angielskim:
 - www.r-bloggers.com
 - www.stackoverflow.com/questions/tagged/r
 - www.stats.stackexchange.com
- Spotkania społeczności R:
 - useR! (od 2004 roku z pominięciem 2005)
 - lokalne konferencje
 - grupy na www.meetup.com

R Users Group Meetup Groups



Groups
165

Members
51,123

Interested
10,305

Cities
133

Countries
34

Literatura

- Kilkadziesiąt książek po angielsku, między innymi seria „Use R!” (Springer)
- Polskie pozycje (m.in. Przewodnik po pakiecie R. Przemysław Biecek, ale jest tego znacznie więcej)
- Czasopisma poświęcone R (Journal of Statistical Software, The R Journal)
- www.rseek.org (wyszukiwarka)

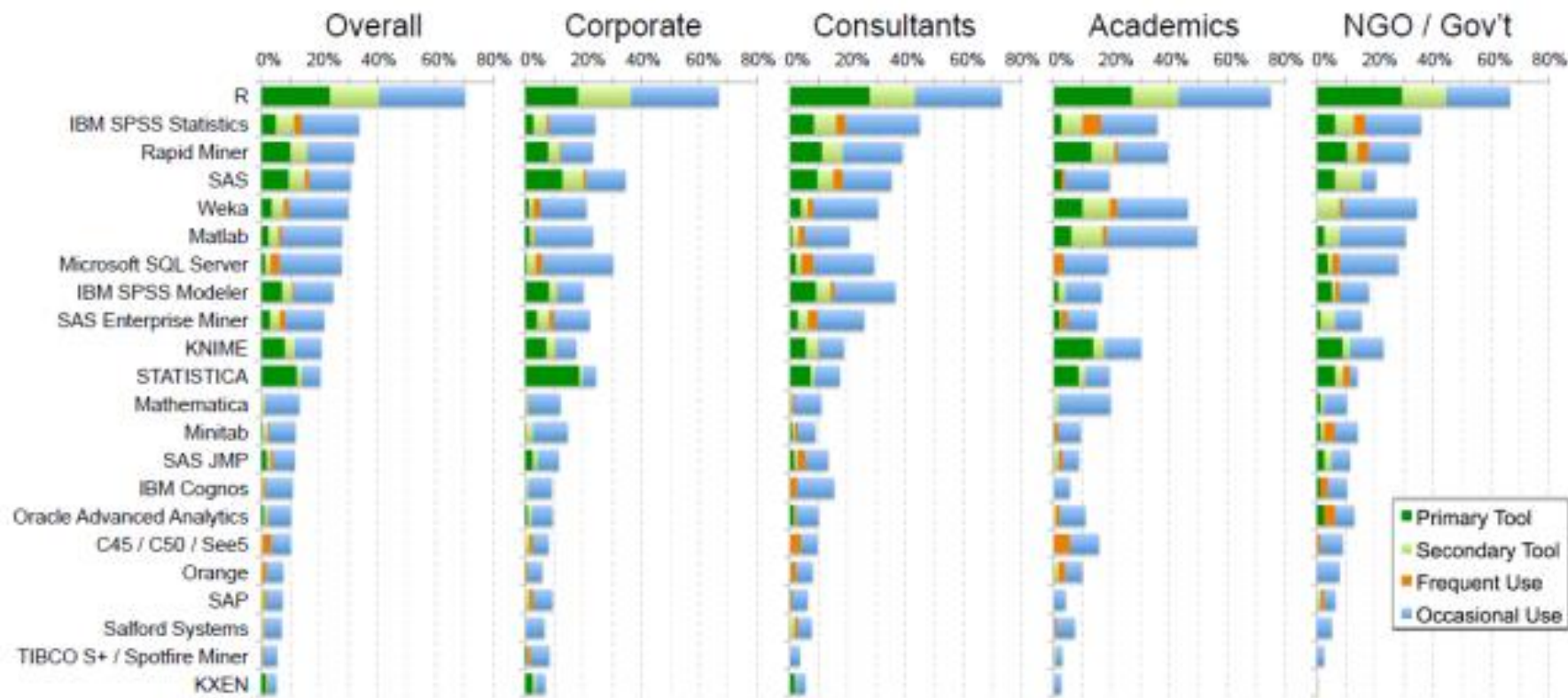
Konkurencyjność

- Atrakcyjność środowiska R spowodowała, że komercyjne programy do analizy danych musiały wyjść naprzeciw i zintegrować się z R:
 - Statistica
 - SAS
 - SPSS
 - Itd...

I nie tylko...

- Sporo innych programów nie będących konkurencją także jest zintegrowanych z R
- R obsługuje wiele formatów plików, a także może komunikować się z wieloma bazami danych (m.in. wszystko z SQL w nazwie)
- R komunikuje się także w większości języków programowania

Najbardziej popularne narzędzie eksploracji danych



DZIEJE SIĘ W POLSCE



?

[19.1451]

Poland 

[51.9194]

Population (n) 38,415,284

Land (sq km) 304,255

Water (sq km) 8,430

Total territory (sq km) 312,685

RANK: 44.

Population weighted score -0.0629

R Foundation members

Benefactors	0
Donors	0
Supporting Institutions	0
Supporting Members	1
Ordinary Members	0

per 1,000 people

R Foundation members	0.00002603
Conference attendees	0.00072888
CRAN downloads	7.38325923
RUG members	0.00000000
R-bloggers visits	0.97010867

useR! attendees

2004 (Vienna, Austria)	1
2006 (Vienna, Austria)	3
2008 (Dortmund, Germany)	3
2009 (Rennes, France)	4
2010 (Gaithersburg, MD, USA)	3
2011 (Coventry, UK)	2
2012 (Nashville, TN, USA)	1
2013 (Albacete, Spain)	11
2014 (Los Angeles, USA)	1

CRAN downloads

All packages	283,630
ggplot2	2,397
devtools	361
knitr	811
RODBC	874
Rcmdr	758

Google Trend searches

Average	12.60%
R	47.00%
CRAN	5.00%
ggplot	0.00%
R graphics	11.00%
R books	0.00%

R User Groups

R meetups (members) 0 (0)

Other statistics

R-bloggers.com visits	37,287
GitHub users	0.20%

Społeczność R w Polsce

- WZUR – Wrocławski/Warszawski Zlot Użytkowników R:
 - 2008, około 30 uczestników (Kraków 2?)
 - 2009, 52 uczestników (Kraków 3?)
 - 2010, 46 uczestników (Kraków 1?)
 - 2011, brak danych
 - 2012, brak danych
- PAZUR – Polski Akademicki Zlot Użytkowników R:
 - 2014, ponad 100 uczestników + 60 w warsztatach (Kraków 2?)

Społeczność R w Polsce

- SER – Spotkania Entuzjastów R (Warszawa):
 - <http://beta.icm.edu.pl/SER/>
 - Spotkania od lutego 2014
 - Aktualnie odbyło się 5 spotkań
 - Organizowane przez dr. hab. inż. Przemysława Biecka
- PAZUR – Poznański Akademicki Złot Użytkowników R:
 - <http://thinking-in-r.blogspot.com/>
 - Spotkania od marca 2012
 - Aktualnie odbyło się 8 spotkań
 - Organizowane przez studenckie koło naukowe

Społeczność R w Polsce

- Kursy R – organizowane przez koła naukowe
- www.smarterpoland.pl (blog fundacji)
- <https://www2.im.uj.edu.pl/gur/> (forum)
- http://www.goldenline.pl/grupy/Komputery_Internet/r/ (forum)

Społeczność R w Polsce

- Kursy R – organizowane przez koła naukowe
- www.smarterpoland.pl (blog fundacji)
- <https://www2.im.uj.edu.pl/gur/> (forum)
- http://www.goldenline.pl/grupy/Komputery_i_internet/r/ (forum)
- KONIEC ???

... a w Krakowie ???

- R na uczelniach – TAK
- R w pracy – ???
 - Narzędzie główne/narzędzie pomocnicze?
 - Codziennie/czasami?

ERKA - PERSPEKTYWY

Cele na najbliższe miesiące

- Cykliczne spotkania – raz na miesiąc:
 - Minimum 10 spotkań w ciągu roku
 - Minimum 10 uczestników
- Prelekcje prowadzone także przez zaproszonych gości (zarówno ze świata nauki jak i świata biznesu – jesteśmy otwarci na propozycje)
- Każda prelekcja (o ile prelegent wyrazi zgodę), będzie nagrywana i kompilowana z prezentacją. Jeżeli będzie taka możliwość to video zostanie uzupełnione o napisy w języku angielskim. Film i inne materiały będą udostępniane przynajmniej na stronie eRki.

Cele na najbliższe miesiące

- Zorganizowanie warsztatów przy eRce:
 - Standardowa forma, różne poziomy trudności (potrzebni chętni do poprowadzenia)
 - Niestandardowa forma:
 - Tutoriale, czyli kompleksowy przegląd metod z danej dziedziny, danego zagadnienia itp.
 - Wspólny cel, czyli opracowanie projektu (jego założeń) i jego realizacja (pro publico bono lub for fun)
 - Inne pomysły także będą brane pod uwagę



Cele na najbliższe miesiące

- W Krakowie też jest R – promocja eRka w Krakowie i poza Krakowem,
- Zdobywanie kontaktów m.in. poprzez wyjazdy na ogólnopolski lub ogólnoswiatowy (o ile znajdą się sponsorzy) zlot użytkowników R,
- Pozyskiwanie sponsorów na bardziej ambitne projekty,
- Planowane są konkursy (raz na 1 lub 2 spotkania) – więcej szczegółów już wkrótce.

Potrzebna pomoc

- Pomoc w organizacji spotkań
- Pomoc przy rozbudowaniu strony www
- Pomoc przy pracach na tematyczną eRką związaną z listami oczekujących na świadczenia zdrowotne – średni czas oczekiwania

Patronat



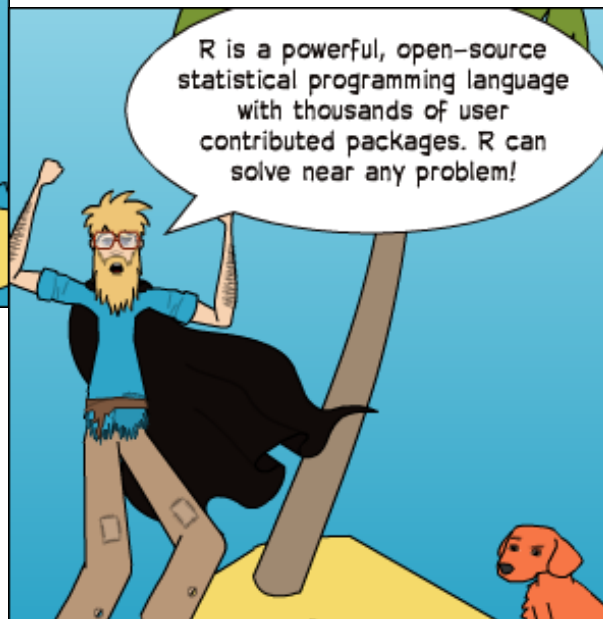
Sponsorzy



eRka #2

10 grudnia o godzinie 18:00

- Miejsce spotkania będzie ustalone w późniejszym terminie
- Jedna z trzech form rejestracji:
 - www.facebook.com (wydarzenie)
 - www.meetup.com (spotkanie)
 - Formularz Google
- Zachęcamy do zgłaszania tematów!



Dziękuję za uwagę

Osoby zainteresowane współpracą lub mające pomysły związane ze społecznością R w Krakowie zapraszam do kontaktu (e-mail: b.sekiewicz@erkakrakow.pl).