

## Online S3 Platform

### Regional scientific production profile

#### 1. Data Collection

##### Overview

We used the website of European Commission to check revealed competitors and regional competitiveness scoreboard in order to select two regions and collect our dataset. Firstly, we selected the province of Utrecht and then selected two major cities, Amersfoort and Zeist. Next step was to find the top 10 competitors from the competitiveness scoreboard, which is based on the degree that firms are active in the same geographical markets (revealed competition). From this scoreboard we selected the region of Dusseldorf which was the strongest competitor of the list.

The data source for this application is [Scopus](#). Scopus is the largest abstract and citation database which was formed to develop an objective system of evaluation and validation of peer-reviewed literature: scientific journals, books and conference proceedings. Everyone can download batches of data containing information on the total number of publications, subject areas with most publications, and year of publications for certain locations which can be useful for creating a benchmarking analysis for countries or cities, for regions, with regard to the performance of each location in relation to science.

#### 2. Subject Area Analysis

##### Table of comparison for scientific fields between two regions

	<i>Dusseldorf-Subject-Area-Analysis.csv</i>	<i>Utrecht-Subject-Area-Analysis.csv</i>
<i>Agricultural and Biological Sciences</i>	1896	9469
<i>Arts and Humanities</i>	652	2827
<i>Biochemistry, Genetics and Molecular Biology</i>	7456	19418
<i>Business, Management and Accounting</i>	467	1328
<i>Chemical Engineering</i>	1301	2132
<i>Chemistry</i>	3673	5429
<i>Computer Science</i>	1195	4110
<i>Decision Sciences</i>	165	576
<i>Dentistry</i>	367	309
<i>Earth and Planetary Sciences</i>	238	7691

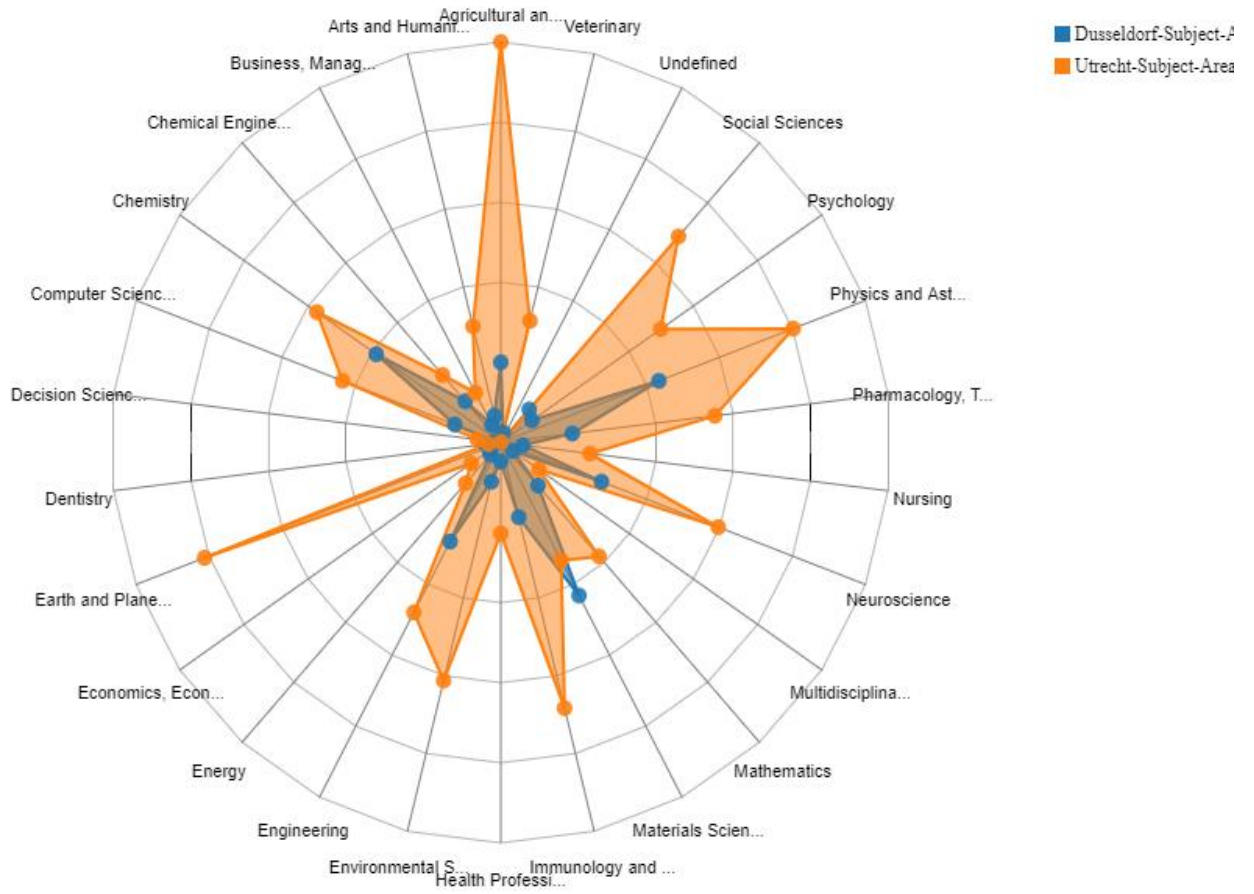
<i>Economics, Econometrics and Finance</i>	332	874
<i>Energy</i>	384	1289
<i>Engineering</i>	2647	4539
<i>Environmental Science</i>	956	5799
<i>Health Professions</i>	452	2152
<i>Immunology and Microbiology</i>	1821	6470
<i>Materials Science</i>	4088	3152
<i>Mathematics</i>	1363	3599
<i>Medicine</i>	17619	50326
<i>Multidisciplinary</i>	355	1125
<i>Neuroscience</i>	2616	5647
<i>Nursing</i>	538	2179
<i>Pharmacology, Toxicology and Pharmaceutics</i>	1750	5235
<i>Physics and Astronomy</i>	4107	7592
<i>Psychology</i>	917	4719
<i>Social Sciences</i>	1042	6507
<i>Undefined</i>	1	9
<i>Veterinary</i>	242	2961

### Comments:

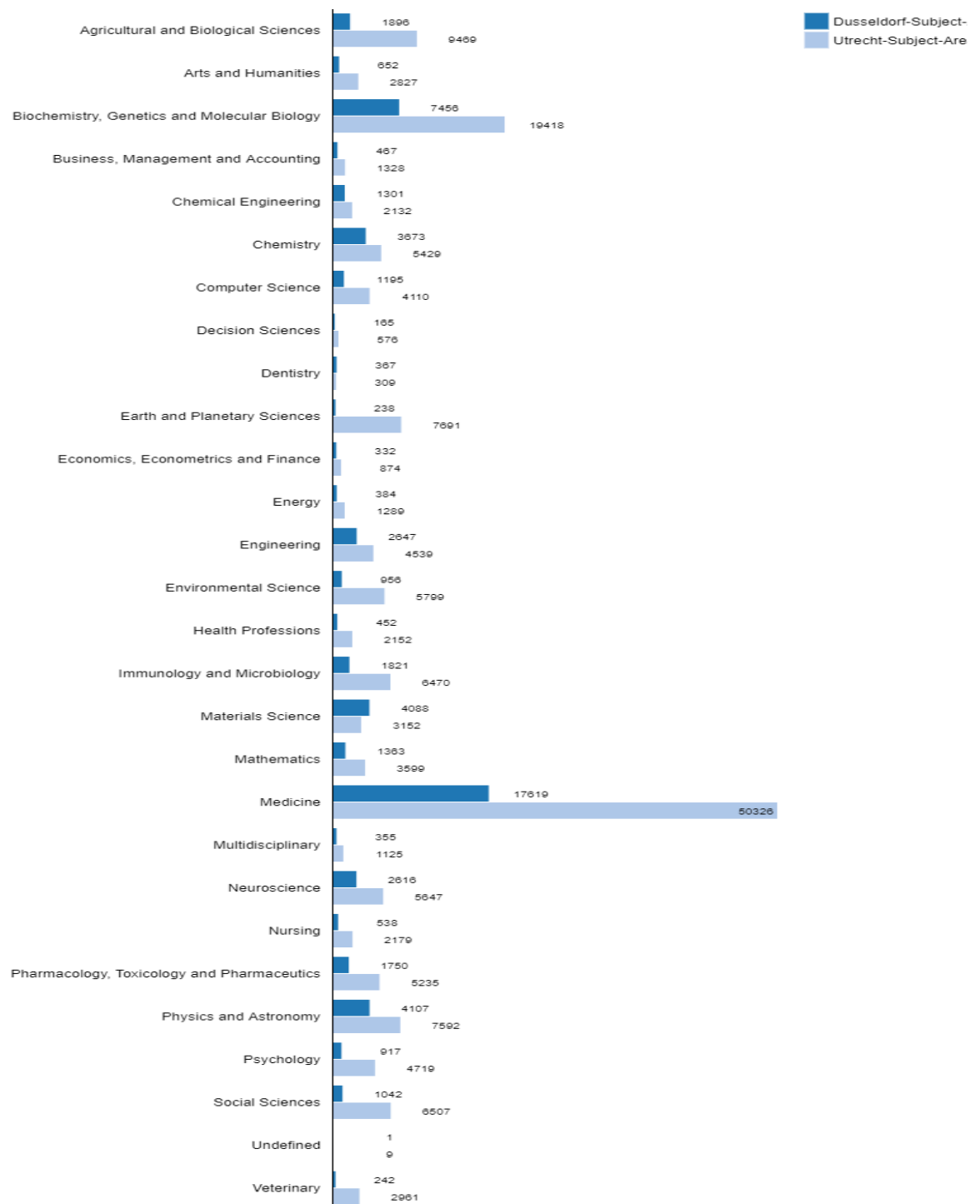
From the first table, we notice some fields between these two regions with the most publications and proceed to comparative conclusions. For instance, number of publications in the field of Medicine is of highest value at both regions, region of Utrecht with 50326 publications and region of Dusseldorf with 17619 publications. Biochemistry Genetics and Molecular Biology are also two field with high value. This analysis helps us to compare all aspects of performance for these regions and to identify common patterns of academic systems. In addition to this, the benchmarking is a valuable tool because we can identify weaknesses and strengths, and link them to overall regional performance. For instance, this table highlights that the field of Dentistry contains the lowest number of publications for both regions, so we can assume it as weakness.

Finally, we present two charts, a Radar chart and a Horizontal bar chart. Radar chart is a graphical method of displaying multivariate data in the form of a two-dimensional chart of three or more quantitative variables represented on axes starting from the same point. In our case, it visualizes the intensity of each scientific field, regarding the number of publications, excluding the two fields with highest value (Medicine and Biochemistry Genetics and Molecular Biology), in order to have better visualization and help users come to useful comparative conclusions. Horizontal bar chart, offers a different way of comparison, presents our categorical data with horizontal rectangular bars, with lengths proportional to the values that they represent, each region has its distinct color and users can come again to useful conclusions, for instance identify scientific fields with highest and lowest number of publications between the two regions.

# Radar Chart



## Horizontal Bar chart



### 3. Affiliation Analysis

	<i>Utrecht-Affiliation-Analysis.csv</i>
<i>Utrecht University</i>	46920
<i>University Medical Center Utrecht</i>	36767
<i>Erasmus University Medical Center</i>	6489
<i>Academic Medical Centre, University of Amsterdam</i>	5928
<i>University of Amsterdam</i>	5085
<i>VU University Medical Center</i>	4607
<i>Nederlandse Organisatie voor toegepast natuurwetenschappelijk onderzoek- TNO</i>	4463
<i>Radboud University Nijmegen Medical Centre</i>	4433
<i>Leiden University Medical Center - LUMC</i>	4056
<i>University of Groningen, University Medical Center Groningen</i>	4033
<i>Wageningen University and Research Centre</i>	3941
<i>Utrecht Institute of Pharmaceutical Sciences</i>	3777
<i>University of Groningen</i>	3709
<i>Rudolf Magnus Institute of Neuroscience</i>	3274
<i>National Institute of Public Health and the Environment</i>	3159
<i>Debye Institute</i>	3146
<i>Vrije Universiteit Amsterdam</i>	2993
<i>Wilhelmina Children's Hospital</i>	2821
<i>Maastricht University</i>	2818
<i>Radboud University Nijmegen</i>	2382
<i>SRON Netherlands Institute for Space Research</i>	2044
<i>Bijvoet Center for Biomolecular Research</i>	2022
<i>Delft University of Technology</i>	2016
<i>Leiden University</i>	1919
<i>NIVEL - The Netherlands Institute for Health Services Research</i>	1852
<i>CBS Fungal Biodiversity Center CBS - KNAW</i>	1779
<i>University of Oxford</i>	1766

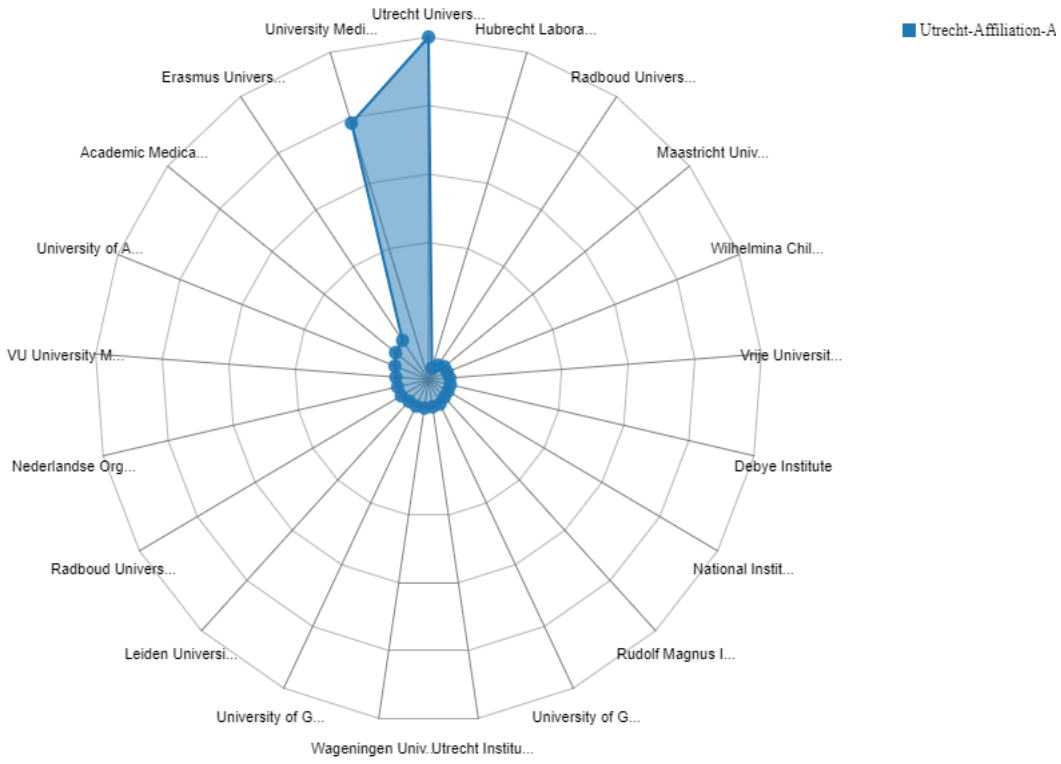
<i>Interuniversity Cardiology Institute of the Netherlands ICIN - KNAW</i>	1741
<i>The Netherlands Cancer Institute</i>	1722
<i>Hubrecht Laboratory KNAW Netherlands Institute for Developmental Biology</i>	1661

Comments:

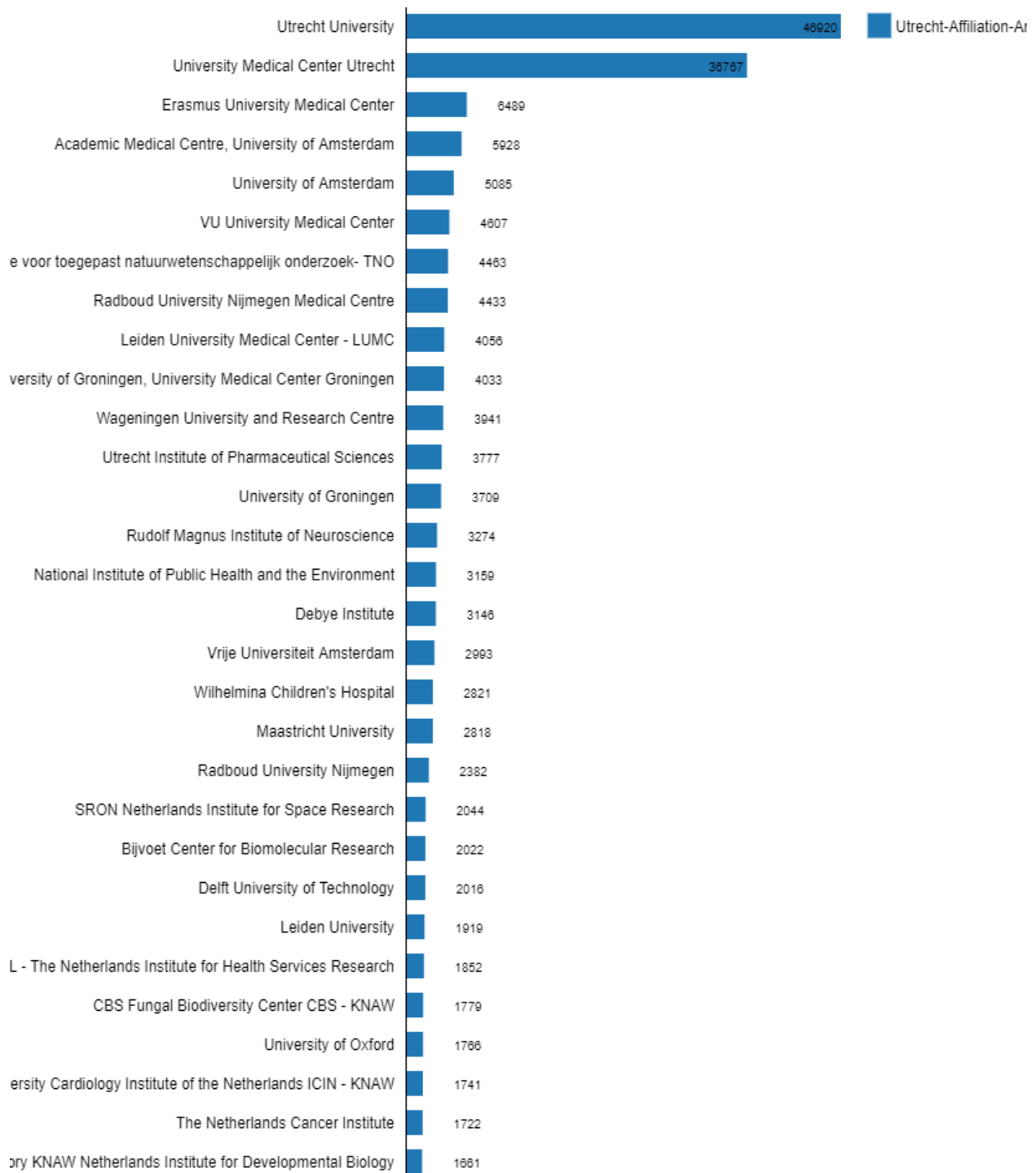
From the first table, we notice two columns, the left one contains names of Institutions, Research centers or Universities and the right one has their values in descending order, presenting the total number of publications in the region of Utrecht. For instance, Utrecht University, University Medical Center of Utrecht and Erasmus University Medical Center, are the top three institutions with the most publications, in the same way we can identify that in the last position are The Netherlands Cancer Institute and Hubrecht Laboratory KNAW Netherlands Institute for Developmental Biology. This analysis helps us to compare all aspects of performance for institutions at the region of Utrecht. In addition to this, the benchmarking is a valuable tool because we can identify weaknesses and strengths, and link them to overall regional performance.

Finally, we present to charts, a Radar chart and a Horizontal bar chart. Radar chart is a graphical method of displaying multivariate data in the form of a two-dimensional chart of three or more quantitative variables represented on axes starting from the same point. In our case, it visualizes the intensity of the top 20 scientific institutions, regarding the number of publications, excluding all other institutions with lower values (such as Debye Institute, SRON Netherlands Institute for Space Research etc.), to have sharper visualization and help users come to useful conclusions. Horizontal bar chart, offers a different way of comparison, presents our categorical data with horizontal rectangular bars, with lengths proportional to the values that they represent. Users can come to conclusions in an easier way and identify at once, institutions with the highest and lowest number of publications in the region of Utrecht.

# Radar Chart



## Horizontal Bar chart





#### 4. Annual Analysis

	<i>Utrecht-Annual-Analysis</i>	<i>Dusseldorf-Annual-Analysis</i>
<i>2000</i>	3425	1355
<i>2001</i>	3370	1415
<i>2002</i>	3698	1477
<i>2003</i>	4252	1715
<i>2004</i>	4228	1640
<i>2005</i>	4766	1844
<i>2006</i>	5111	1893
<i>2007</i>	5188	1976
<i>2008</i>	5502	2003
<i>2009</i>	5795	2002
<i>2010</i>	6028	2037
<i>2011</i>	6660	2175
<i>2012</i>	7291	2402
<i>2013</i>	7596	2440
<i>2014</i>	7649	2592
<i>2015</i>	7762	2750
<i>2016</i>	7950	2702
<i>2017</i>	6278	2096
<i>2018</i>	10	8

#### Comments:

From the first table, we notice three columns, the first one contains the timeframe between 2000 and 2018, the second and third presents the total number of publications per year, for Utrecht's and Dusseldorf's region accordingly. For instance, the year with highest value at Utrecht, was 2016 with 7950 publications, but for the region of Dusseldorf, value 2750 is the highest number of publications at year 2015. In the same way we can easily identify that in the last position is the year 2018 because the database at Scopus hasn't been update yet. This analysis helps us to understand the increased rate of growth in scientific publications every year. In addition to this, the benchmarking is a valuable tool because we can identify weaknesses and strengths, and link them to yearly regional performance.

Finally, we present a line chart (time series chart), which displays information and shows a trend in our dataset, over intervals of time. It contains two axes, the horizontal axis presents the timeframe, and the vertical axis presents the number of publications. The chart visualizes the increased rate of growth in publications regarding time evolution, with two colorful lines that are drawn chronologically.

## Time Series Chart

