1. How many unique pages did you find? Uniqueness is established by the URL, but discarding the fragment part. So, for example, *http://www.ics.uci.edu#aaa* and *http://www.ics.uci.edu#bbb* are the same URL.

141 / 47159

1. What is the longest page in terms of number of words? (HTML markup doesn’t count as words)

11948 words

1. What are the 50 most common words in the entire set of pages? (**Ignore English stop words**, which can be found, for example, [here Links to an external site.](https://www.ranks.nl/stopwords)) Submit the list of common words ordered by frequency.

Research 21842

Student 11441

2022 10844

Students 10240

Undergraduate 10135

will 10102

Graduate 10079

BS 10031

Software 10030

2018 9466

2020 9259

The 9251

2017 9224

2021 9189

2019 9175

September 9144

Computer 9114

Engineering 9099

2016 8946

2015 8945

Courses 8841

News 8803

Ramesh 8484

Information 8480

We 8477

Past 8149

Books 8007

PhD 7771

Informatics 7757

Department 7693

Policies 7533

Current 7367

October 7329

design 6909

Projects 6907

June 6827

July 6776

Design 6524

data 6477

November 6394

May 6302

Science 6218

March 6132

Course 6124

April 6102

January 6085

December 6052

February 6027

UCI 5989

Profiles 5988

1. How many subdomains did you find in the ics.uci.edu domain? Submit the list of subdomains ordered alphabetically and the number of unique pages detected in each subdomain. The content of this list should be lines containing *URL***,***number,*for example:  
   http://vision.ics.uci.edu, 10 (not the actual number here)

132 subdomains

accessibility.ics.uci.edu: 1

acoi.ics.uci.edu: 62

aiclub.ics.uci.edu: 1

archive-beta.ics.uci.edu: 2

archive.ics.uci.edu: 3

asterix.ics.uci.edu: 7

asterixdb.ics.uci.edu: 4

auge.ics.uci.edu: 2

calendar.ics.uci.edu: 1

cbcl.ics.uci.edu: 4

cert.ics.uci.edu: 2

cgvw.ics.uci.edu: 1

checkmate.ics.uci.edu: 1

chenli.ics.uci.edu: 9

cherry.ics.uci.edu: 1

chime.ics.uci.edu: 1

cloudberry.ics.uci.edu: 2

cml.ics.uci.edu: 32

cocoa-krispies.ics.uci.edu: 1

code.ics.uci.edu: 18

codeexchange.ics.uci.edu: 1

computableplant.ics.uci.edu: 34

coronavirustwittermap.ics.uci.edu: 1

cradl.ics.uci.edu: 39

create.ics.uci.edu: 10

cwicsocal18.ics.uci.edu: 19

cyberclub.ics.uci.edu: 2

dataguard.ics.uci.edu: 1

dataprotector.ics.uci.edu: 1

dblp.ics.uci.edu: 2

dejavu.ics.uci.edu: 1

dgillen.ics.uci.edu: 16

duttgroup.ics.uci.edu: 51

dynamo.ics.uci.edu: 1

edgelab.ics.uci.edu: 17

elms.ics.uci.edu: 1

emj-pc.ics.uci.edu: 1

emj.ics.uci.edu: 77

esl.ics.uci.edu: 2

evoke.ics.uci.edu: 15

flamingo.ics.uci.edu: 11

fr.ics.uci.edu: 4

futurehealth.ics.uci.edu: 5

grape.ics.uci.edu: 91

graphics.ics.uci.edu: 11

graphmod.ics.uci.edu: 2

hack.ics.uci.edu: 1

hai.ics.uci.edu: 3

hana.ics.uci.edu: 4

hobbes.ics.uci.edu: 1

hombao.ics.uci.edu: 1

honors.ics.uci.edu: 1

hpi.ics.uci.edu: 8

i-sensorium.ics.uci.edu: 1

iasl.ics.uci.edu: 27

ics.uci.edu: 8

informatics.ics.uci.edu: 28

intranet.ics.uci.edu: 1

ipf.ics.uci.edu: 7

ipubmed.ics.uci.edu: 6

ironwood.ics.uci.edu: 1

isg.ics.uci.edu: 284

jgarcia.ics.uci.edu: 26

jujube.ics.uci.edu: 2

kdd.ics.uci.edu: 1

keys.ics.uci.edu: 1

linguistics.uci.edu: 1

luci.ics.uci.edu: 10

mailman.ics.uci.edu: 3

malek.ics.uci.edu: 1

map125.ics.uci.edu: 1

mapgrid.ics.uci.edu: 1

mcs.ics.uci.edu: 118

mdogucu.ics.uci.edu: 1

mds.ics.uci.edu: 18

mhcid.ics.uci.edu: 23

mlearn.ics.uci.edu: 1

mondego.ics.uci.edu: 6

mse.ics.uci.edu: 1

mswe.ics.uci.edu: 19

nalini.ics.uci.edu: 12

ngs.ics.uci.edu: 3167

omni.ics.uci.edu: 1

password.ics.uci.edu: 1

pasteur.ics.uci.edu: 2

perennialpolycultures.ics.uci.edu: 1

physics.uci.edu: 6

plrg.ics.uci.edu: 22

psearch.ics.uci.edu: 2

redmiles.ics.uci.edu: 11

riscit.ics.uci.edu: 1

sconce.ics.uci.edu: 1

sdcl.ics.uci.edu: 93

se.ics.uci.edu: 1

seal.ics.uci.edu: 6

seraja.ics.uci.edu: 1

sherlock.ics.uci.edu: 2

sli.ics.uci.edu: 16

sourcerer.ics.uci.edu: 1

sprout.ics.uci.edu: 2

stairs.ics.uci.edu: 1

statconsulting.ics.uci.edu: 5

student-council.ics.uci.edu: 2

studentcouncil.ics.uci.edu: 4

support.ics.uci.edu: 2

swiki.ics.uci.edu: 2

tad.ics.uci.edu: 1

tastier.ics.uci.edu: 1

tippers.ics.uci.edu: 1

tippersweb.ics.uci.edu: 4

transformativeplay.ics.uci.edu: 61

tutors.ics.uci.edu: 1

ugradforms.ics.uci.edu: 1

unite.ics.uci.edu: 15

vision.ics.uci.edu: 15

wearablegames.ics.uci.edu: 15

wics.ics.uci.edu: 1237

www-db.ics.uci.edu: 25

www-db.ics.uci.edu:8000: 1

www-db.ics.uci.edu:8003: 1

www-db.ics.uci.edu:8004: 1

www-db.ics.uci.edu:8006: 1

www.cert.ics.uci.edu: 1

www.economics.uci.edu: 9

www.graphics.ics.uci.edu: 1

www.ics.uci.edu: 1105

www.informatics.ics.uci.edu: 1

www.informatics.uci.edu: 1774

www.isg.ics.uci.edu: 1

www.physics.uci.edu: 113

www.statistics.uci.edu: 1

xtune.ics.uci.edu: 1