

# StatCrunch -laskentasovellus

## Yleistä sovelluksesta

StatCrunch on Integrated Analytics LLC:n valmistama sovellus tilastotieteellisten analyysien tuottamista varten. Se on verkon yli käytettävä analyysisovellus, jonka tarkoitus on korvata kokonaan perinteinen työpöytäsovellus. Siihen kuuluu joukko valmiita laskentatoimintoja ja kyky tuottaa erilaisia graafisia esityksiä.

StatCrunchin toiminta perustuu verkkoon, eli sitä käytetään www-selaimen avulla. StatCrunch on toteutettu Java-applettina, joten sen toiminta ei riipu selaimen versiosta. Lähestulkoon kaikista nykyisistä verkkoselaimista löytyy tuki Java-appleteille. Selainsovelluksissa on joitakin etuja perinteisiin sovelluksiin nähden. Sovellus pitää asentaa vain yhdelle palvelimelle, jos ei se ole käytettävissä verkossa jo valmiiksi. Verkon yli sovellusta voidaan käyttää mistä tahansa, missä verkkoyhteys on saatavilla. Sen käyttöliittymä on myös samanlainen käyttöympäristöstä riippumatta.

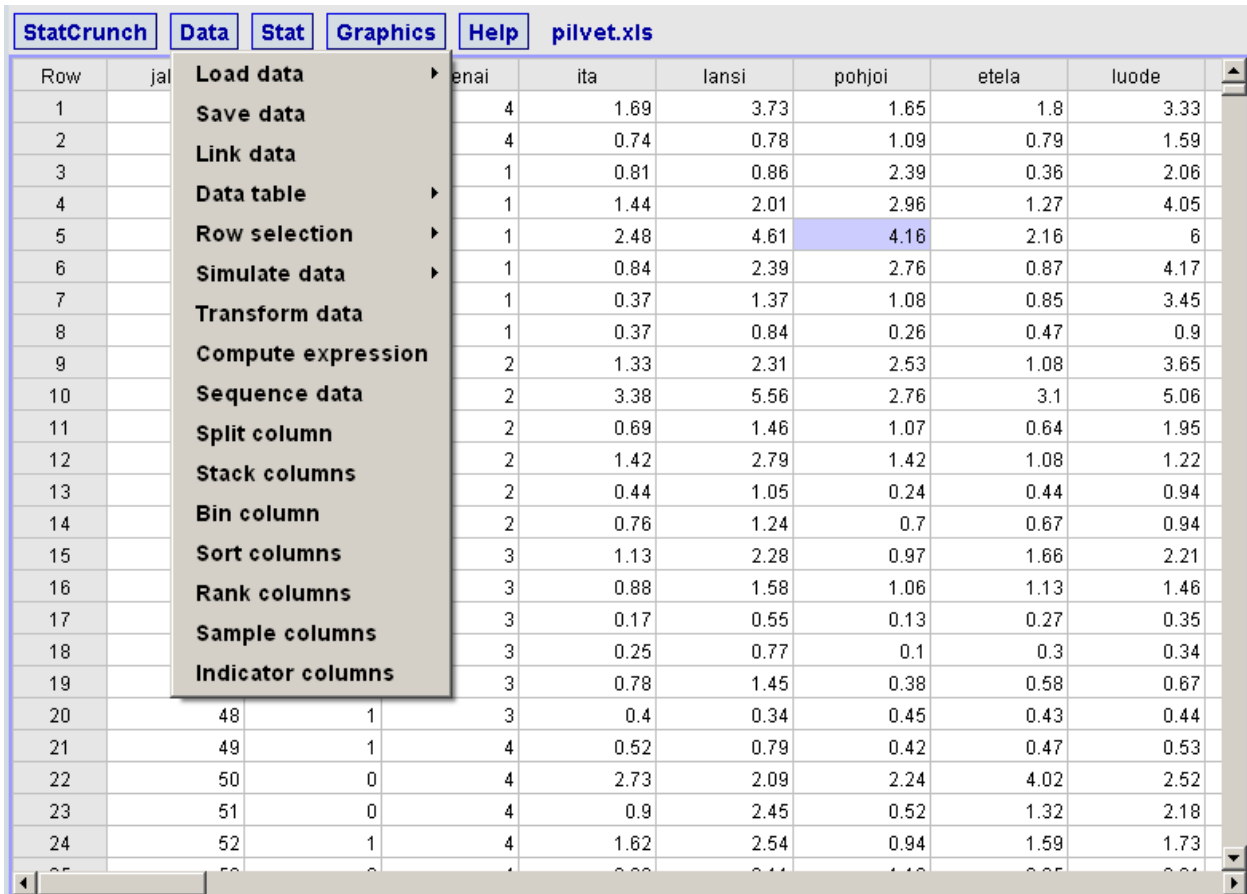
Ohjelmistoa on mahdollista käyttää suoraan valmistajan sivulta. Tämä vaatii lisenssin ostamisen; opiskelijalisenssi on tavallista edullisempi. Toinen mahdollisuus on hankkia lisenssi ohjelmiston asentamiseen omaan palvelinympäristöön. Ohjelmiston voi vapaasti ladata valmistajan sivustolta, mutta se ei toimi ilman lisenssiä. Tässäkin tapauksessa opiskelukäyttö on hinnoiteltu erikseen. Hintatiedot eivät kuitenkaan ole julkisia.

Testikäyttöä varten StatCrunchista jaetaan kokeilulisenssejä. Ne ovat maksuttomia ja toimivat kymmenen päivää. Varsinaisen käytön aloittaminen ei vaadi käyttäjältä muuta, kuin palvelinsivun avaamisen ja lisenssikoodin syöttämisen sovellukselle.

StatCrunchin käyttäjät muodostavat jonkinasteisen yhteisön, jonka kautta voi halutessaan jakaa muille valmiita datakokonaisuuksia ja raportteja. Molempia on saatavilla valmistajan sivuilla tuhansia.

## Sovelluksen käyttäminen ja toiminnot

Sovellus on käyttövalmis käyttäjän avattua sivuston, jolla sovellus sijaitsee. Käyttöliittymästä on esitetty esimerkki kuvassa 3.1. Sovellukseen on esimerkin vuoksi ladattu dataa kurssin materiaalisivulta.



The screenshot shows the StatCrunch web application interface. At the top, there are menu tabs: StatCrunch, Data, Stat, Graphics, and Help. The current file is named 'piilvet.xls'. A data table is displayed with columns: Row, jal, enai, ita, lansi, pohjoi, etela, and luode. The 'Data' menu is open, showing various options like 'Load data', 'Save data', 'Link data', 'Data table', 'Row selection', 'Simulate data', 'Transform data', 'Compute expression', 'Sequence data', 'Split column', 'Stack columns', 'Bin column', 'Sort columns', 'Rank columns', 'Sample columns', and 'Indicator columns'. The table data is as follows:

Row	jal	enai	ita	lansi	pohjoi	etela	luode
1		4	1.69	3.73	1.65	1.8	3.33
2		4	0.74	0.78	1.09	0.79	1.59
3		1	0.81	0.86	2.39	0.36	2.06
4		1	1.44	2.01	2.96	1.27	4.05
5		1	2.48	4.61	4.16	2.16	6
6		1	0.84	2.39	2.76	0.87	4.17
7		1	0.37	1.37	1.08	0.85	3.45
8		1	0.37	0.84	0.26	0.47	0.9
9		2	1.33	2.31	2.53	1.08	3.65
10		2	3.38	5.56	2.76	3.1	5.06
11		2	0.69	1.46	1.07	0.64	1.95
12		2	1.42	2.79	1.42	1.08	1.22
13		2	0.44	1.05	0.24	0.44	0.94
14		2	0.76	1.24	0.7	0.67	0.94
15		3	1.13	2.28	0.97	1.66	2.21
16		3	0.88	1.58	1.06	1.13	1.46
17		3	0.17	0.55	0.13	0.27	0.35
18		3	0.25	0.77	0.1	0.3	0.34
19		3	0.78	1.45	0.38	0.58	0.67
20	48	1	3	0.4	0.34	0.45	0.43
21	49	1	4	0.52	0.79	0.42	0.47
22	50	0	4	2.73	2.09	2.24	4.02
23	51	0	4	0.9	2.45	0.52	1.32
24	52	1	4	1.62	2.54	0.94	1.59
25	53	0	4	0.88	0.44	1.16	0.84

*StatCrunch-sovelluksen käyttöliittymä.*

Suurin osa näyttötilasta on oletusarvoisesti varattu taulukolle, jossa esitetään tutkittavaa dataa. Tätä tilaa voidaan käyttää myös tulosten esittämiseen. Ikkunan yläreunasta löytyvät valikot, joista sovelluksen toimintoja käytetään. Kuvassa on näistä eräs avattuna. Ikkunan yläreunassa esitetään myös ladatun datatiedoston nimi. Seuraavaksi käydään läpi sovelluksen valikot ja niihin liittyvä toiminnallisuus pääpiirteissään.

## *StatCrunch*

Ensimmäinen valikko on saman niminen kuin itse sovellus. Siitä voidaan valita ainoastaan, missä analysointitulokset esitetään. Valinta tapahtuu kohdasta `Session results`. Tulosten esittämiseen voidaan käyttää osaa ruudusta, koko ruutua tai uutta ikkunaa.

## *Data*

Toinen valikko käsittelee nimensä mukaisesti enimmäkseen datan lataamista ja organisointia sovellukseen. Dataa voi syöttää sovellukseen manuaalisesti. Sitä voidaan ladata verkko-osoitteista, leikepöydältä tai useista erilaisista tiedostoista kohdasta `Load data`. Esimerkiksi Excel-muotoiset tiedostot soveltuvat käytettäväksi sellaisenaan. Samasta valikosta voi myös kohdasta `Save data` tallettaa dataa eri muodoissa tietokoneelle, jolla työskennellään.

Valikosta löytyy useita komentoja datan manipulointiin ja järjestelyyn halutulla tavalla. Arvoille voi suorittaa erilaisia laskutoimituksia niin halutessaan, esimerkiksi kohdasta `Transform data` voidaan muuttujia muun muassa luokitteluasteikollisia muuttujia luokitella uudelleen. Tarkasteluun voi myös valita haluamiaan osajoukkoja tutkittavasta datasta. StatCrunch osaa myös kohdasta `Simulate data` itse tuottaa esimerkkidataa kolmeentoista erilaiseen jakaumaan perustuen, aina Bernoullin jakaumasta Weibull-jakaumaan.

## *Stat*

StatCrunchin tärkeimmät analyysityökalut löytyvät Stat-valikosta. Erilaisten perustietojen kokoaminen tilastomuuttujista onnistuu helposti; muutaman painalluksen takaa kohdasta `Summary Stats` saadaan aikaan yhteenveto halutuista muuttujista. Se sisältää erilaisia tunnuslukuja, kuten minimi, maksimi, keskiarvo, mediaani ja keskihajonta. Kvantitatiivisten muuttujien tutkimiseen on komentoja enemmänkin, kohdasta `Summary Stats` saadaan selville myös muuttujien välinen korrelaatiokerroin.

Ristiintaulukointi tapahtuu StatCrunchissa kohdasta `Tables`. Tämä on toteutettu sovelluksessa samalla tavalla kuin useimmat monimutkaiset käskyt. Sovellus kyselee

tarvittavat parametrit, kuten käytettävät muuttujat, käyttäjältä niin kutsutun velhotoiminnon avulla. Annettuja arvoja voi muuttaa vapaasti kunnes lopuksi suoritetaan tulosten laskenta. Menetelmä on varsin helppokäyttöinen ja itse asiassa melko samankaltainen esimerkiksi SPSS Statistics -ohjelmiston vastaavan kanssa. Kohdasta `Tables` onnistuu myös taulukon muuttujien välinen riippuvuustarkastelu ja frekvenssijakauman tuottaminen.

T-testi tehdään kohdasta `T statistics`, samoin kuin luottamusvälin laskenta. Oleellimmat kurssilla tarvittavat analyysimenetelmät ovat siis löydettävissä StatCrunchista varsin helposti. Esimerkiksi tunnuslukujen tuottaminen tapahtuu todella vaivattomasti.

### *Graphics*

StatCrunchin neljäs valikko keskittyy erilaisten graafisten esitysten tuottamiseen. Nämä ovat sovelluksen vahva osa-alue: mahdollisia esityksiä on paljon ja erilaisten esitysten ulkoasu voidaan yksilöidä kohdasta `Color Schemes`. Kuten analyysien, myös graafisten esitysten tuottaminen tapahtuu selkeän velhotoiminnon avulla.

Tilastomuuttujien tarkasteluun soveltuvia esityksiä ovat esimerkiksi pylväsdiagrammi, joka löytyy kohdasta `Bar Plot`, ympyrädiagrammi kohdasta `Pie Chart` tai histogrammi kohdasta `Histogram`. Pisteparven luominen tapahtuu kohdasta `Dotplot`. Kaikkien muuttujien riippuvuutta toisistaan voi havainnollistaa sirontakaavioilla kohdasta `Pairs Plot`.

Graafiset esitykset voidaan tallettaa tiedostona tietokoneelle, jolla työskennellään. Niitä voi myös kopioida leikepöydälle tai tulostaa suoraan sovelluksesta.

### *Help*

Viimeisestä valikosta löytyvät sovelluksen ohjeet. Käyttöohjeet ovat englanniksi, kuten muukin sovellus. Kaikille tärkeimmille toiminnoille löytyy verraten kattava ohjeistus, mutta hakutoimintoa ei ole.

## **Yhteenveto**

Oleellisimmat kurssilla tarvittavat analyysimenetelmät ovat löydettävissä StatCrunchista varsin helposti. Esimerkiksi tunnuslukujen tuottaminen tapahtuu todella vaivattomasti. Graafisten esitysten luonti on monipuolista ja kurssin puitteisiin varsin riittävää. Ohjelmiston yleiskäyttöisyydellä on merkitystä myös sen käytön kannalta kurssin ulkopuolella. StatCrunchissa käytetyt menetelmät ovat samoja kuin muissakin alan sovelluksissa, joskin niiden käyttötapa voi vaihdella.

StatCrunchin käyttöliittymä on yksinkertainen ja intuitiivinen. Peruskäyttöön ohjelma soveltuu hyvin; aineiston lataaminen, tallettaminen ja tunnuslukujen tarkastelu on vaivatonta. Sen käytön opetteluun kynnys on matala, mutta luonnollisesti sen laskentamahdollisuudet eivät ole yhtä monipuoliset kuin esimerkiksi SPSS Statistics -sovelluksen vastaavat.

### **Lisätietoa:**

<http://www.statcrunch.com/>