

# Statistiek I voor de Bedrijfseconomische Wetenschappen Studiewijzer 2024–2025

Luc Hens  
Vrije Universiteit Brussel

## Contactinformatie

**Titularis:** Luc Hens (luc.hens@vub.be). Pleinlaan 5, 5de verdieping, lokaal 515 (PL5 5.15). Als je mij of één van de assistenten een mail stuurt met betrekking tot deze cursus, zet als onderwerp “Statistiek I Je.voornaam Je.achternaam.”

**Assistenten:** Nathalie Colpin (nathalie.colpin@vub.be, lokaal PL5 520); Julie Orlans (julie.inge.a.orlans@vub.be, room PL5 521).

**Spreekuren:** Als je vragen hebt over de cursus, wend je dan eerst tot je assistent tijdens, voor, of na de werkcolleges. Maak eventueel een afspraak. Als de assistent je niet kan helpen, ben je welkom tijdens de pauze van de hoorcolleges of tijdens mijn spreekuur (vrijdag 14:00–16:45: maak een afspraak op [calendly.com/luc-hens](https://calendly.com/luc-hens)).

**Leerplatform:** Materiaal met betrekking tot de cursus (deze studiewijzer, de datasets, het formuleblad, en meer) staat op de cursuspagina van het Canvas-leerplatform dat je kan bereiken via het studentenportaal.

**Hoorcolleges (HOC):** Maandag 16:00–18:00, Aula QA

**Werkcolleges (WPO):** raadpleeg je uurrooster.

## Inhoud

Deze cursus bestrijkt beschrijvende statistiek (hoe gegevens weergeven en samenvatten) en introduceert de student tot de inferentiële statistiek (hoe gegevens uit een steekproef gebruiken om numerieke uitspraken te doen over een populatie). We leren hoe we best gegevens weergeven in tabellen en grafieken (frequentietabellen, kruistabellen, histogrammen, boxplots, tijdreeksdiagrammen, ...); hoe we de vorm van een empirische verdeling kunnen beschrijven en de centrale tendentie en de spreiding ervan kunnen samenvatten (gemiddelde, mediaan, standaarddeviatie, interkwartielafstand, ...); en hoe we het verband tussen twee variabelen grafisch kunnen weergeven, numeriek

beschrijven en interpreteren (spreidingsdiagram, correlatie, regressielijn). We maken dan kennis met de basiselementen van de kansrekening die nodig zijn voor de inferentiële statistiek (regels van de kansrekening, gezamenlijke en voorwaardelijke kans, contingentietabellen, kansbomen,...); en met kansvariabelen (verwachte waarde en standaarddeviatie; kansmodellen). Tenslotte bestuderen we de steekproevenverdelingen van fracties en gemiddelden, en gebruiken de Centrale-Limietstelling om betrouwbaarheidsintervallen te berekenen voor populatiefracties en -gemiddelden

## Voorkennis

De wiskundige voorkennis die je nodig hebt voor deze cursus is op middelbare-schoolniveau (richtingen uit het secundair onderwijs met minstens 4 uur wiskunde per week). Het gaat meer bepaald om de volgende delen uit het handboek van de voorbereidingscursus wiskunde (Carette et al., 2013): hoofdstuk 2: Rekeningtechnieken; hoofdstuk 6: Relaties (enkel 6.5 Rechten en 6.9 Exponentiële en logaritmische functies); en hoofdstuk 8: Vergelijkingen en Ongelijkheden (enkel 8.1.1 Lineaire vergelijkingen in één onbekende, en 8.5.1 Ongelijkheden in één onbekende). In de kalender van de colleges op het einde van deze studiewijzer verwijs ik naar de relevante delen uit Carette et al. (2013).

## Leerdoelen

Het doel van deze cursus is om inzicht te geven in de beschrijvende statistiek (gegevens weergeven en beschrijven) en de student in te leiden in de inferentiële statistiek (geldige veralgemeningen maken op basis van steekproefgegevens). Op het einde van deze cursus ben je in staat om:

- kwalitatieve en kwantitatieve variabelen te onderscheiden, weer te geven en te beschrijven;
- een correlatiecoëfficiënt en regressielijn te berekenen en te interpreteren;
- de regels van de kansrekening toe te passen;
- te werken met kansvariabelen en kansmodellen (binomiaal model, normaal model, ...);
- uit te leggen welke eigenschappen de steekproevenverdeling van een fractie of een gemiddelde heeft, en aan te geven onder welke voorwaarden deze eigenschappen geldig zijn;
- een betrouwbaarheidsinterval voor een fractie of een gemiddelde te berekenen (als aan de voorwaarden is voldaan), en de betekenis van een betrouwbaarheidsinterval te interpreteren;
- uit te leggen wat de beperkingen zijn van statistische methoden;
- op een ethische manier om te gaan met statistiek;

- statistische software en een wetenschappelijke rekenmachine te gebruiken om statistische berekeningen te doen (gegevens ingeven; beschrijvende statistieken te berekenen en grafieken maken; de kans te berekenen die bij een normaal verdeelde variabele hoort; betrouwbaarheidsintervallen berekenen);
- de resultaten van statistisch werk te communiceren, en meer bepaald om de resultaten van een statistische analyse weer te geven in een geschreven verslag dat bestaat uit een niet-technische samenvatting voor de beslissers (zodat zij betere beslissingen kunnen nemen), en een technisch deel gericht naar mensen met dezelfde expertise als jezelf dat de technische details en exacte interpretatie van de resultaten weergeeft. Het verslag is opgemaakt in *APA Style*.

## Studiemateriaal

**Handboek** Het verplichte handboek is Sharpe et al. (2021). Koop de paperback Global Edition (ISBN 978-1-292-26931-3, ongeveer 70 euro). Die is aanzienlijk goedkoper dan de Amerikaanse editie. Verkrijgbaar bij de Standaard Student Shop (gebouw B, begane grond op de hoek) en online (<https://www.amazon.com.be/-/en/Norean-Sharpe/dp/1292269316/>, <https://www.bol.com/be/nl/p/sharpe-bsnsstats-ge-p4/9300000018478480/>). Als je een vorige editie gebruikt is het je eigen verantwoordelijkheid om na te gaan wat de verschillen met de meest recente editie zijn. Appendix B bevat een correspondentietabel tussen de vorige en de huidige editie.

**Rekenmachine** Je hebt voor de werkcolleges en het examen een Texas Instruments **TI-84** rekenmachine nodig. De aanbevolen versie is de TI-84 Plus CE-T maar andere versies van de TI-84 voldoen ook. Je zal de TI-84 ook gebruiken in Statistiek II. De statistische functies van de TI-84 die we in Statistiek I gebruiken staan op het formuleblad. Hier zijn enkele webwinkels waar je de TI-84 Plus CE-T kan kopen (ongeveer 140 euro): <https://www.coolblue.be/nl/product/865206/texas-instruments-ti-84-ce-t-python-app.html>; <https://www.amazon.de/-/en/Instruments-Python-Graphic-Calculator-Display/dp/B08F3PBLCL>; <https://www.dreamland.be/e/nl/dl/texas-instruments-rekenmachine-ti-84-plus-ce-t-python-972868>. Hou de bon en de originele verpakking bij. Dat maakt het makkelijker om later je rekenmachine door te verkopen. Mocht je al een **TI-NSpire** hebben, kan je die ook gebruiken. Je moet dan zelf uitzoeken hoe je de statistische functies kan oproepen.

**Statistische software** Een rekenbladprogramma is nuttig om eenvoudige berekeningen te doen of om eenvoudige tabellen en lijsten te bewerken, maar niet om ernstig statistisch werk te verrichten. Statistici en datawetenschappers gebruiken gespecialiseerde statistische software zoals R, SPSS, of Stata, of een programmeertaal zoals Python. In deze cursus geef ik een inleiding tot het gebruik van de statistische programmeeromgeving R. R is openbron, gratis, en is samen met Python de meest gebruikte programmeeromgeving voor statistiek en *data science*. Je zal R verder gebruiken in Statistiek II, Econometrie, en

je onderzoeksprojecten. Bijlage A legt uit hoe je R en RStudio Desktop (een *integrated development environment (IDE)* voor R) kan afhalen en installeren op je eigen computer.

**Gegevensbestanden** Omdat elke statistisch programma, rekenbladprogramma of tekstbewerkingsprogramma op elke computer kan werken met tekstbestanden, zijn tekstbestanden de geprefereerde opmaak voor gegevensbestanden. De **gegevensbestanden** die in het handboek worden gebruikt staan hier: <https://das1.datadescription.com/datafiles/>. Ze staan ook als een comprimeerd zip-bestand op Canvas dat je kan downloaden en dan openen door er op te dubbelklikken.

## Bijkomend studiemateriaal

De beste manier om statistiek te leren is door (veel) oefeningen te maken. De antwoorden op de oneven oefeningen bij elk hoofdstuk staan achteraan in Sharpe et al. (2021). Je kan nog meer opgeloste oefeningen vinden in Kazmier (2004) (of een meer recente editie). Beschikbaar in de VUB-bibliotheek, plaatsnummer 311.17 G KAZM 2004. Dalgaard (2008) is een goede gids tot R. Beschikbaar in de VUB-bibliotheek, plaatsnummer 004.9 G DALG 2008. Naast R en de TI-84 is WolframAlpha ([wolframalpha.com](http://wolframalpha.com)) een handig instrument om statistische berekeningen te doen. Dit document legt uit hoe je de statistische berekeningen uit deze cursus kan doen in WolframAlpha: <https://luc-hens.github.io/statistics-in-wolfram-alpha.html>.

## Wijze van examineren

Datum, tijd en plaats van het examen worden aangekondigd op de webstek van de faculteit. Het examen is schriftelijk en bestaat geheel of gedeeltelijk uit meerkeuzevragen. Bij de quoterings van meerkeuzevragen corrigeren we voor gissen met de methode van hogere cesuur; dat betekent bijvoorbeeld dat je van 20 meerkeuzevragen met vier antwoordmogelijkheden, er minstens 12 correct moet beantwoorden om te slagen.

Voor heel wat van de meerkeuzevragen moet je, om het juiste antwoord te vinden, de oplossing uitvoerig uitschrijven zoals we je tijdens de werkcolleges hebben aangeleerd. De vragenbundel heeft voldoende ruimte om dat te doen op de achterkant van het vorige blad. Schrijf eerst de gegevens op, dan wat gevraagd is. Schrijf dan de oplossing uit, stap voor stap. Leid elke belangrijke stappen in met een korte verklarende zin. Je doet voor jezelf (niet voor ons: we lezen dit niet bij het verbeteren). De uitwerking is echter essentieel om het juiste antwoord te vinden en je werk op het einde van het examen nog eens grondig na te lezen.

Breng het volgende mee naar het examen: je studentenkaart, je gelamineerde formuleblad (als het formuleblad niet gelamineerd is mag je het niet gebruiken), een blauwe of zwarte balpen, een vulpotlood (met zachte vullingen: HB nr. 2), enkele kleurpotloden (niet rood), een gom, een meetlat met centimeterschaal, en de TI-84. Zorg er voor dat de rekenmachine volledig is opgeladen (TI-84 CE-T) of dat je reservebatterijen mee hebt (TI-84 Plus). Wis het geheugen van

de TI-84 vóór je naar het examen komt. Stop alles (behalve het formuleblad) in een doorzichtig hersluitbaar plastic zakje (Ziploc, 1 liter). Zet je smartphone op stil en berg smartphone en smartwatch op in je boekentas of rugzak. **Geen** pennenzak, smartphone, smartwatch, eigen papier, of voedsel op je werkplek. Een flesje water mag als je het etiket verwijdt (herbruikbare flesjes zijn beter; je kan een Dopperfles met het VUB-logo kopen in de Standaard Student Shop). Neem je plastic zakje met inhoud (inclusief de TI-84) en je gelamineerde formuleblad en ga in stilte naar de door ons aangegeven plek. Berg jas en tas buiten bereik op onder je bank. Eens je zit, wacht je tot iedereen een plek heeft.

## Voorbereiding, aanwezigheid

Aanwezigheid is cruciaal: statistiek is een moeilijk onderwerp, en studenten die meerdere hoor- of werkcolleges missen slagen zelden voor deze cursus. Zorg dat je, voor je naar de hoorcolleges en de werkcolleges komt het hoofdstuk uit Sharpe et al. (2015) dat we gaan bestrijken hebt gelezen. Tijdens de werkcolleges los je, begeleid door een assistent, oefeningen op die aan het einde van elk hoofdstuk staan. Aanwezigheid en actieve deelname tijdens de werkcolleges is verplicht.

Breng het volgende mee naar **hoor- en werkcolleges**: A4-papier met vierkante ruitjes (5 mm), een vulpotlood (met zachte vullingen: HB nr. 2), enkele kleurpotloden, een gom, en een meetlat met centimeterschaal. Breng daarenboven ook mee naar de **werkcolleges** (WPO's): je handboek (want daarin staan de opgaves van de oefeningen die je tijdens het werkcollege maakt); de TI-84 CE-T; het gelamineerde formuleblad; en je laptop met RStudio geïnstalleerd en de gegevensbestanden afgehaald. Het gebruik van laptops, tablets of smartphones tijdens hoorcolleges is niet toegestaan, en tijdens werkcolleges mag je enkel een laptop gebruiken voor oefeningen waar dat nodig is: onderzoekers hebben wetenschappelijk aangetoond dat ze het jezelf en je burens moeilijker maken om iets op te steken van de colleges (Sana et al., 2013). Maak tijdens de hoorcolleges notities.

## Kalender van de colleges

Lees het aangeduide materiaal *vóór* je naar de colleges komt. De hoofdstukken verwijzen naar Sharpe et al. (2021). Ik beveel sterk aan om extra oefeningen op te lossen, individueel of in groep. De antwoorden op de oneven oefeningen bij elk hoofdstuk staan achteraan in Sharpe et al. (2021). De assistent kan je helpen als je vast zit, maar breng dan wel je geschreven voorbereiding mee.

### Week 22 (begint op maandag 10 februari 2025)

**Hoe deze cursus is georganiseerd.** Lees: Maxine Rappé (2024, 28 november). Weg met laptops? Notities nemen doe je beter met de hand, zegt Noorse studie. VRTNWS (<https://www.vrt.be/vrtnws/nl/2024/11/27/laptop-schrijven-hand-papier-hersenen-leren-studie/>). Doe vóór het eerste werkcollege het volgende: Koop de Texas Instruments **TI-84 Plus CE-T** rekenmachine. Haal R en RStudio af en installeer op je computer. Haal de gegevensbestanden (als .txt-bestanden) van het handboek af van Canvas. Haal het formuleblad af van het leerplatform. Ga naar een copierzaak en laat het formuleblad afdrukken

op dik papier. Schrijf je naam in drukletters rechtsboven. Laat dan het formuleblad lamineren. Je mag het formuleblad gebruiken op het schriftelijk examen, maar enkel als het gelamineerd is.

**Module 1: Gegevens en beslissingen.** Hoofdstuk 1. Oefeningen 2, 4, 6, 12, 16, **19, 23, 25**.

**Module 2: Kwalitatieve gegevens weergeven en beschrijven.** Hoofdstuk 2. Oefeningen: 2, 4, 6, 8, 10, 28, 40, **46**.

**Werkcollege:** Breng je laptopcomputer (met daarop R en RStudio geïnstalleerd) mee naar het werkcollege: je zal leren werken met R en RStudio. Zorg er voor dat je laptopcomputer volledig is opgeladen.

### **Week 23 (begint op maandag 17 februari 2025)**

**Module 2: Kwalitatieve gegevens weergeven en beschrijven.** Hoofdstuk 2 (vervolg).

### **Week 24 (begint op maandag 24 februari 2025)**

**Module 3: Kwantitatieve gegevens weergeven en beschrijven.** Hoofdstuk 3. Sla “Transforming skewed data” (pp. 111–113) over. In het college zal ik een derde manier introduceren om gegevens in een histogram weer te geven, nl. een densiteitshistogram. Oefeningen: 2, 4, 6, 8, 12, 14, 18, 23, **30**.

### **Week 25 (begint op maandag 3 maart 2025)**

**Module 3: Kwantitatieve gegevens weergeven en beschrijven** (vervolg). Oefeningen **42, 48**. Voor deze oefeningen heb je RStudio nodig: breng je laptop mee naar het werkcollege. Zorg dat je R en RStudio hebt geïnstalleerd en de gegevensbestanden hebt afgehaald.

### **Week 26 (begint op maandag 10 maart 2025)**

**Module 4: Correlatie en Regressie.** Hoofdstuk 4. Sla delen 4.7, 4.8, 4.10, 4.11 over. Oefeningen: **2**, 4, 6, 9, 14 (eerst zonder de statistische functies van de TI-84 te gebruiken; controleer dan je resultaten aan de hand van de `LinReg`-functie van de TI-84), **19, 20, 24, 25, 30, 32, 60, 70** (gebruik de `LinReg`-functie op de TI-84 en RStudio: vergeet dus niet om je laptop mee te brengen en de gegevensbestanden af te halen).

### **Week 27 (begint op maandag 17 maart 2025)**

**Module 5: Kansrekening.** Hoofdstuk 5. Sla deel 5.9 (“Bayes’ rule”) over. Oefeningen: 2, 4, 6, 8, 10, **12, 14, 56, 58**.

### **Week 28 (begint op maandag 24 maart 2025)**

**Module 5: Kansrekening** (vervolg).

## Week 29 (begint op maandag 31 maart 2025)

**Module 6: Toevalsvariabelen en kansmodellen.** Hoofdstuk 6. Sla “The Geometric Model” (p. 234) en “The Poisson Model” (pp. 238–239) over. Oefeningen: 1, 2, 4, 6, 10, 14, **16**, 18, 26, 42.

## Week 30 (begint op maandag 7 april 2025)

Lentevakantie. Geen les.

## Week 31 (begint op maandag 14 april 2025)

Lentevakantie. Geen les.

## Week 32 (begint op maandag 21 april 2025)

Maandag 21 april 2025 is een wettelijke feestdag. Geen les.

**Module 6: Toevalsvariabelen en kansmodellen (in de werkcolleges, vervolg)** Oefeningen: 1, 2, 4, 6, 10, 14, **16**, 18, 26, 42.

## Week 33 (begint op maandag 28 april 2025)

Donderdag 1 mei 2025 is een wettelijke feestdag.

**Module 7: De normale verdeling.** Hoofdstuk 7. Sla deel 7.6 over. Lees wél het kaderstukje “What can go wrong” onderaan p. 272—dat hoort niet bij deel 7.6. Oefeningen: **2**, 4, 6, **10**, 12, 18, **28** (met de 68-95-99,7-regel), **30** (met de 68-95-99,7-regel), 32, 34, 36, **40**, 42, 60.

## Week 34 (begint op maandag 5 mei 2025)

**Module 8: Observationele studies, enquêtes, experimenten.** Hoofdstuk 8; hoofdstuk 9 (enkel delen 9.1 en 9.2, pp. 312-318, 335) (ik zal deze pagina’s beschikbaar maken op Canvas voor wie de derde editie gebruikt). Bekijk de videoclip “Leading questions,” een fragment uit de sitcom *Yes Prime Minister*, reeks 1 episode 2 (“The Ministerial Broadcast”), BBC, 1986 (<https://youtu.be/G0ZZJXw4MTA>) (video en transcriptie van de dialoog). Oefeningen voor hoofdstuk 8: 2, 4, 12, 16, 28, 30, 48. Oefeningen voor hoofdstuk 9: **2**, **4**. **Huiswerk:** gebruik voor elke oefening van module 7 (zie vorige week) die de TI-84 vereiste RStudio om de berekeningen te doen. Druk het Console-paneel met de resultaten af en breng mee naar het werkcollege.

## Week 35 (begint op maandag 12 mei 2025)

**Module 9: Steekproevenverdeling van een fractie en betrouwbaarheidsinterval van een fractie.** Hoofdstuk 10. Oefeningen: 2, 4, 8, 10, 12; 11 & 14 (die samen horen), 16, 18 (gebruik in deel *b* 99,7% in plaats van 99%), 48, 50, **58**. Bereken voor alle oefeningen waarbij je een betrouwbaarheidsinterval moet vinden eerst het betrouwbaarheidsinterval zonder de 1-PropZInt-functie

van de TI-84 te gebruiken; verifieer dan of je (bijna) hetzelfde resultaat krijgt met de 1-PropZInt-functie.

## Week 36 (begint op maandag 19 mei 2025)

**Module 10: Steekproevenverdeling van een gemiddelde en betrouwbaarheidsinterval van een gemiddelde als de steekproef groot is.** Hoofdstuk 11. Enkel delen 11.1 tot 11.3, en: “Cautions about Interpreting Confidence Intervals” (p. 393). We beperken ons in Statistiek I tot gevallen waar de steekproef voldoende groot is om de steekproevenverdeling van een gemiddelde te benaderen door de Normale verdeling (delen 11.1 tot 11.3). In het kaderstukje van deel 11.5 (Sharpe et al., 2021, bovenaan p. 386) kan je dan  $t_{n-1}^*$  vervangen door de kritische waarde  $z^*$  uit het Normale model. In dergelijke gevallen is het 95%-betrouwbaarheidsinterval voor een gemiddelde ongeveer

$$\bar{y} \pm 2 \times SE(\bar{y})$$

Oefeningen: **4, 6, 8 & 10** (veronderstel dat je de normale benadering mag gebruiken), **26**.

## Week 37 (begint op maandag 26 mei 2025)

Blokweek.

## Weken 38–42: Examenperiode

Datum, tijd en plaats van het schriftelijk examen worden aangekondigd op de webstek van de faculteit. Het examen bestrijkt al het materiaal vermeld in de bovenstaande kalender (ook als we het niet tijdens de colleges hebben besproken). Lees aandachtig de uitleg over het examen in de sectie “Wijze van examineren” hierboven.

## Bijlage A R en RStudio Desktop installeren

### R and RStudio Desktop installeren (Windows)

Open in je bladerprogramma (Microsoft Edge, Firefox) de downloadpagina van *Posit*, het bedrijf dat RStudio maakt:

<https://posit.co/download/rstudio-desktop/>

Links zie je: “1: Install R.” Klik op de knop “DOWNLOAD AND INSTALL R.” Je bladerprogramma opent een nieuw tabblad met de downloadpagina van R (*The Comprehensive R Archive Network, CRAN*). Selecteer “Download R for Windows.” Klik op “base.” Klik op “Download R 4.2.2 for Windows” (het versienummer kan recenter zijn dan 4.2.2). Je bladerprogramma begint met het afhalen van het bestand R-4.2.2-win.exe (het versienummer kan recenter zijn) en zet het op je computer, waarschijnlijk in de Downloads-map. Eens het bestand op je computer staat, dubbelklik er op om de software te installeren. Het installatieprogramma zet het programma (R.app) in je Programma’s-map.

Ga nu in je bladerprogramma terug naar het tabblad RStudio Desktop:



<https://posit.co/download/rstudio-desktop/>

Rechts zie je “2. Install RStudio.” Klik op de knop “DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS.” Je bladerprogramma begint met het afhalen van het bestand RStudio-2022.12.0-353.exe (het versienummer kan recenter zijn dan 2022.12.0-353) en zet het op je computer (waarschijnlijk in de Downloads-map). Eens het bestand op je computer staat, dubbelklik op het icoontje om de software te installeren.

## R and RStudio Desktop installeren (macOS)

Controleer eerst welke chip je Mac heeft (oudere Macs hebben een Intel-chip, recente Macs hebben een Apple M-chip). Je doet dat als volgt: Ga in de menubalk naar het appeltje (linksboven) en kies: About This Mac > Chip. Noteer of je Mac een Intel- of Apple M-chip heeft. Je hebt die informatie zo meteen nodig.

Open in je bladerprogramma de downloadpagina van *Posit*, het bedrijf dat RStudio maakt:

<https://posit.co/download/rstudio-desktop/>

Links zie je: “1: Install R.” Klik op de knop “DOWNLOAD AND INSTALL R.” Je bladerprogramma opent een nieuw tabblad met de downloadpagina van R (*The Comprehensive R Archive Network, CRAN*). Selecteer “Download R for macOS.” Als je Mac een Apple M-chip heeft, klik op “R-4.2.2-arm64.pkg” (het versienummer kan recenter zijn dan 4.2.2); als je Mac een Intel-chip heeft, klik op “R-4.2.2.pkg” (het versienummer kan recenter zijn dan 4.2.2). Je bladerprogramma begint met het afhalen van het bestand en zet het op je computer, waarschijnlijk in de Downloads-map. Eens het bestand op je computer staat, dubbelklik er op om de software te installeren. Het installatieprogramma zet het programma (R.app) in je Programma’s-map.

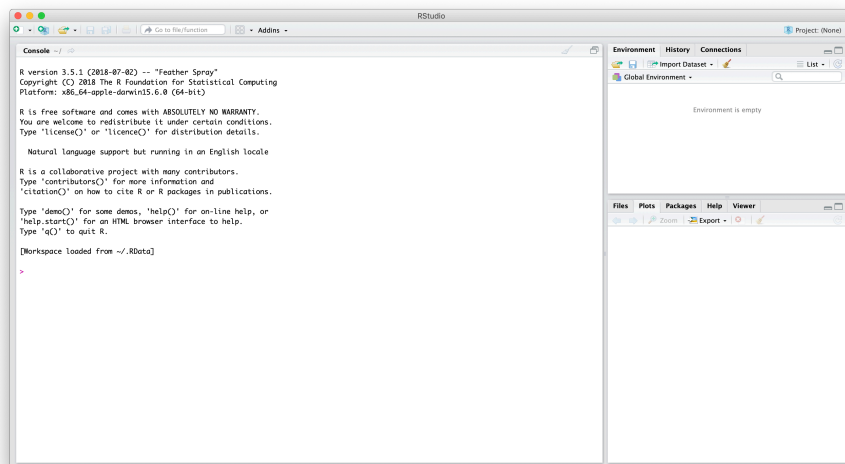
Ga nu in je bladerprogramma terug naar het tabblad RStudio Desktop:

<https://posit.co/download/rstudio-desktop/>

Rechts zie je “2. Install RStudio.” Klik op de knop “DOWNLOAD RSTUDIO DESKTOP FOR MAC.” Je bladerprogramma begint met het afhalen van het bestand RStudio-2022.12.0-353.dmg (het versienummer kan recenter zijn dan 2022.12.0-353) en zet het op je computer (waarschijnlijk in de Downloads-map). Eens het bestand op je computer staat, dubbelklik op het icoontje om de software te installeren. Een venster “RStudio-Desktop-2022.12.0-353” gaat open (het versienummer kan recenter zijn). Aan de linkerzijde van dat venster zie je een icoontje van je Programma’s-map; aan de rechterzijde is een icoontje met de naam RStudio.app. Sleep het RStudio.app-icoontje (rechts) naar de Programma’s-map (links) en laat los: het installatieprogramma zet RStudio.app in de Programma’s-map. Ga naar de Programma’s-map en sleep het RStudio-icoontje naar je Dock; zo kan je er altijd makkelijk aan. Ga nu naar de Finder en open een nieuw Finder-venster (Archief > Nieuw Finder-Venster). Ga in het linkse paneel naar beneden tot je onder “Locaties” ziet staan: RStudio-2022.12.0-353 (het versienummer kan recenter zijn). Klik op de uitwerp-knop (*eject*).

## R laten draaien in RStudio

Nadat je R en RStudio hebt geïnstalleerd op je PC of Mac, localiseer het RStudio-icoontje in je Programma's-map of Dock in macOS, of waar je programma's zichtbaar zijn in Windows. Klip op het RStudio-icoontje. RStudio start en toont een venster zoals in figuur 1. Het paneel aan de linkerzijde heet



Figuur 1: RStudio-venster (in macOS)

Console. In het Console-paneel verschijnt tekst: het versienummer van R en informatie over het programma, gevolgd door een prompt (`>`) en een knipperende cursor die aangeeft dat R wacht op wat je ingeeft. Typ na de prompt:

```
1 + 2
```

en druk op de return-toets. Het Console-paneel toont je het resultaat:

```
[1] 3
```

Het resultaat wordt voorafgegaan door `[1]` om aan te geven dat dit de eerste regel is van het resultaat—sommige resultaten lopen over meerdere regels.

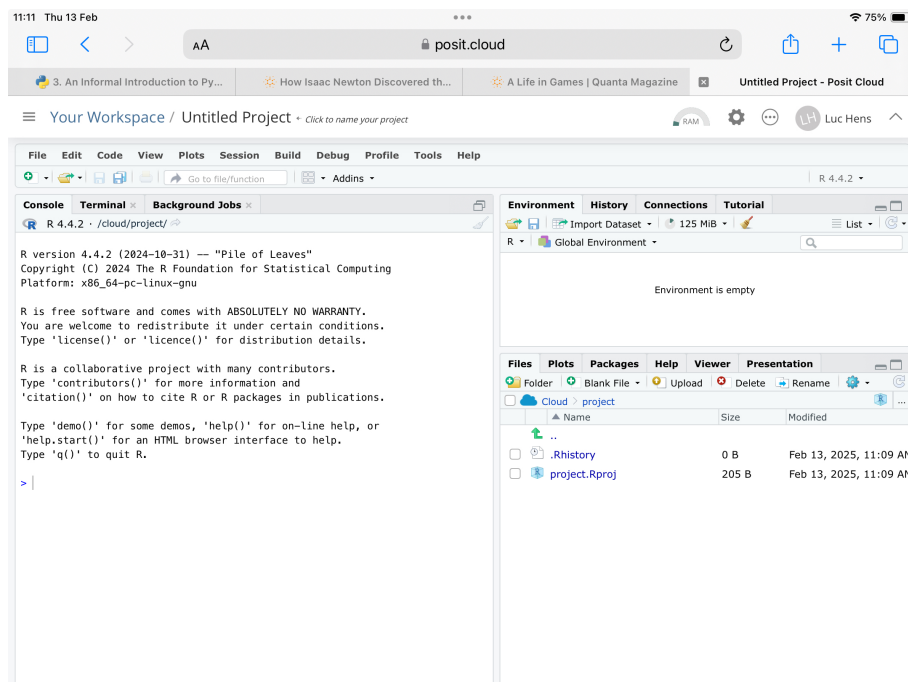
Je bent nu klaar om berekeningen te doen! Je kan de `.dmg`-, `.pkg`-, of `.exe`-bestanden die in je Downloads-map zitten in de prullenmand gooien.

## R and RStudio on a tablet or Chromebook

If you don't have a computer running macOS, Windows or Linux, you can still run R and RStudio on a tablet (an iPad or a tablet with an Android-based operating system) or a Chromebook. This also works to run R and RStudio in a browser window on a Chromebook or on a computer without having installed R and RStudio on the computer.

Download the R scripts and data files that you need to your tablet. Files you download will typically end up in the Downloads folder in the Files or My Files app of the tablet operating system. For more information for iPadOS see the support article on “Find files and folders in Files on iPad,” <https://support.apple.com/en-gb/guide/ipad/ipadb3b759ed/ipados>; for Android see <https://www.androidpolice.com/find-downloads-android-phone-tablet/>.

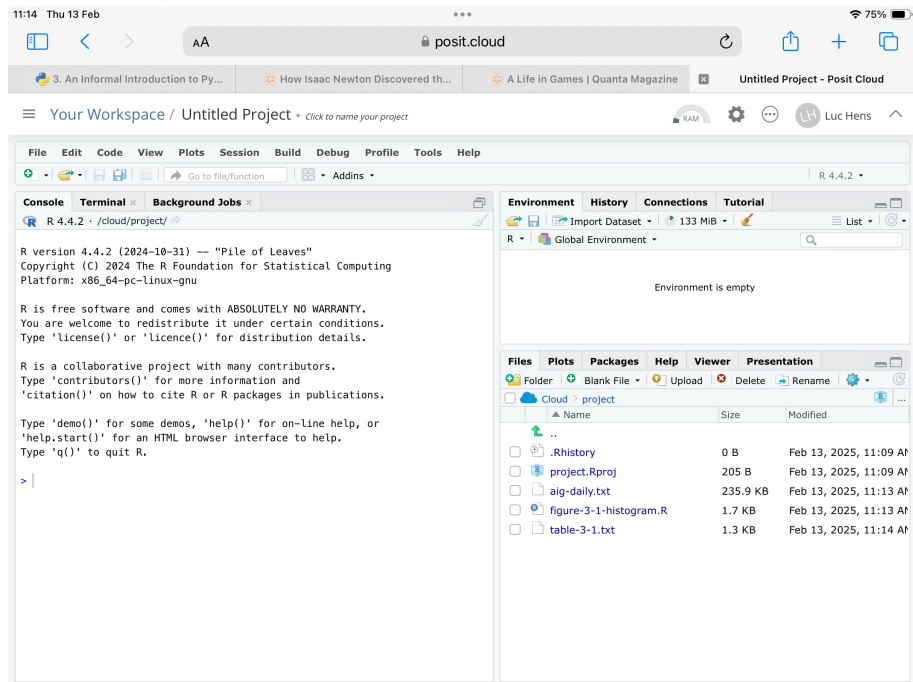
Open the browser app on your tablet and sign up with **posit Cloud** (<https://posit.cloud>). If you are a student, the free version probably will do. Then log in to **posit Cloud**. Create a new RStudio project. You get a window similar to the one in figure 2.



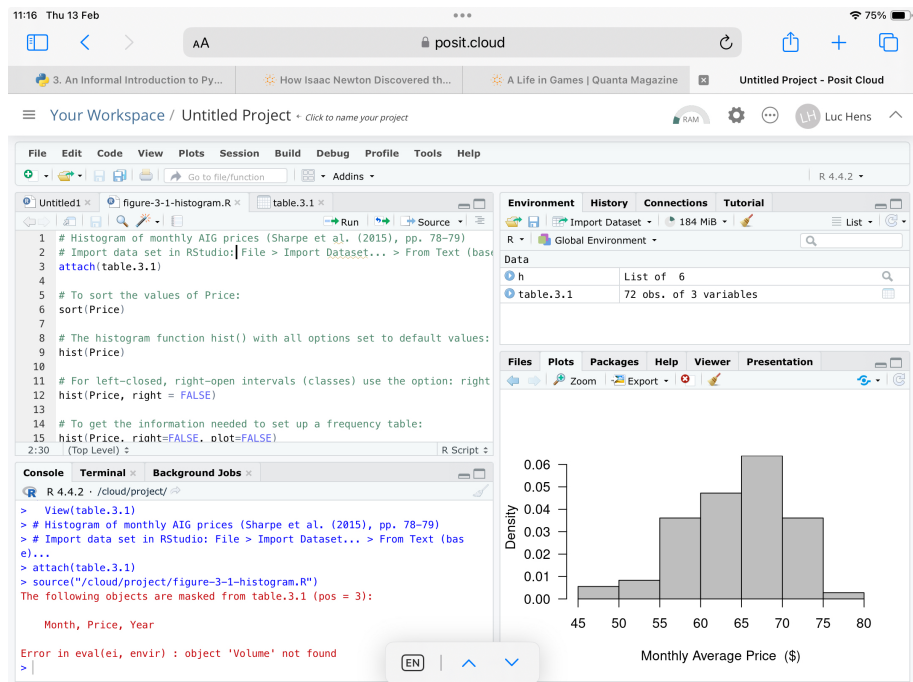
Figur 2: RStudio window in posit.Cloud at startup

Now upload the R scripts and data files you need to the RStudio environment on your tablet. You do so by pressing the “Upload” button in the bottom right panel and navigating to the location on your tablet where you stored the files, typically the Downloads folder. The files will show up in bottom right panel. I uploaded the data files aig.daily.txt and table-3-1.txt and the script figure-3-1-histogram.R (see figure 3).

You can now open the R script and import the data files that are listed in the bottom right panel. To open a script, in the top left panel press the + button below File. Select R Script. Select the Folder icon to the right of the + button. Select the R script from the pop-up window (I selected figure-3-1-histogram.R). Then import the data file: in the top left panel select File > Import Dataset > From Text (base)... , and select the file in the pop-up window. Then select the lines from the script you want to run and press the Run button. The result will look like figure 4.



Figur 3: RStudio window in posit.Cloud after uploading data and scripts



Figur 4: RStudio window in posit.Cloud after running the script

## Bijlage B Correspondentie tussen Sharpe et al. (2015, 2021)

| Sharpe et al. (2015) | Sharpe et al. (2021)                                   |
|----------------------|--------------------------------------------------------|
| Hoofdstuk 1          | Hoofdstuk 1                                            |
| Hoofdstuk 2          | Hoofdstuk 2                                            |
| Hoofdstuk 3          | Hoofdstuk 3                                            |
| Hoofdstuk 4          | Hoofdstuk 4                                            |
| Hoofdstuk 5          | Hoofdstuk 5                                            |
| Hoofdstuk 6          | Hoofdstuk 6                                            |
| Hoofdstuk 7          | Hoofdstuk 7                                            |
| Hoofdstuk 8          | Hoofdstuk 8 + nieuw Hoofdstuk 9 ( <i>Experiments</i> ) |
| Hoofdstuk 9          | Hoofdstuk 10                                           |
| Hoofdstuk 11         | Hoofdstuk 11                                           |

De oefeningen die in de kalender van de colleges zijn opgegeven voor Sharpe et al. (2021) komen grotendeels overeen met Sharpe et al. (2015) (soms verschilt de nummering of zijn de gegevens bijgewerkt). Oefeningen die nieuw zijn in Sharpe et al. (2021) zijn aangegeven in **vetjes**.

## Bijlage C Errata in Sharpe et al. (2021)

**Ontbrekende eenheden in tabellen.** De eenheden (\$) voor de variabele **Price** ontbreken in tabel 1.1 op p. 37. De waarden in kolom 4 moeten zijn: \$5.99, \$9.99, \$9.99, \$10.99, \$11.99. Alternatief kan je de eenheden aangeven in de hoofding van de tabel: **Price (\$)**.

**Ontbrekende eenheden in berekeningen.** Neem steeds de **eenheden** mee in je berekeningen zodat je weet in welke eenheden het resultaat van je berekeningen is uitgedrukt. Zo is de correcte berekening van de  $z$ -score voor de woning met een waarde van \$340.000 op p. 101:

$$z = \frac{\$340.000 - \$175.000}{\$55.000} = 3,00$$

Je kan nu zien dat de eenheden in de teller en in de noemer hetzelfde zijn (namelijk \$), en dat het resultaat van de deling dus geen eenheden heeft. Ook in het voorbeeld bovenaan p. 150 ontbreken de eenheden. Er moet staan:

$$\hat{\text{Price}} \approx \$13\,439 + 113.12 \text{ \$/sq.ft} \times (\text{Living Area in sq. ft})$$

Bovenaan p. 151 moet dan staan:

$$\hat{\text{Price}} \approx \$13\,439 + 113.12 \text{ \$/sq.ft} \times 3000 \text{ sq.ft} \approx \$352\,799$$

Merk op dat de eenheden van het resultaat van de berekeningen nu volgen uit de berekening. Bovendien is de interpretatie van een coëfficiënt is veel helderder als je de eenheden ervan kent.

**Afrondingen.** Gebruik  $\approx$  als je afrondt of een benadering gebruikt, niet  $=$ . Zie het voorbeeld in de vorige paragraaf.

Ook elders laten Sharpe et al. (2021) de eenheden verkeerdelijk weg in de berekeningen, en gebruiken ze verkeerdelijk een gelijkheidsteken ( $=$ ) als er een  $\approx$  moet staan.

## Referenties

- Carette, P., Guerry, M.-A., Theuns, P., and Vanderhoeft, C. (2013). *Brugcursus Wiskunde voor Humane Wetenschappen*. VUB Press, Brussel.
- Dalgaard, P. (2008). *Introductory Statistics with R*. Springer, Berlin, 2<sup>nd</sup> edition.
- Kazmier, L. J. (2004). *Schaum's Outline of Theory and Problems of Business Statistics*. Schaum's Outline Series. McGraw-Hill, New York, 4th edition.
- Sana, F., Weston, T., and Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education*, 62:24–31.
- Sharpe, N. R., De Veaux, R., and Velleman, P. (2015). *Business Statistics*. Pearson Education, 3<sup>rd</sup> edition.
- Sharpe, N. R., De Veaux, R., and Velleman, P. (2021). *Business Statistics*. Pearson Education, 4<sup>th</sup> edition.