

## Contact

1433 North LeLand Ave  
Indianapolis, IN 46219  
3174537784 (Mobile)  
harlananelson@gmail.com

[www.linkedin.com/in/harlananelson](https://www.linkedin.com/in/harlananelson)  
(LinkedIn)  
[stackoverflow.com/users/4629916/harlan-nelson](https://stackoverflow.com/users/4629916/harlan-nelson) (Personal)  
[github.com/harlananelson](https://github.com/harlananelson)  
(Personal)  
[www.wyzant.com/match/tutor/79745010](https://www.wyzant.com/match/tutor/79745010) (Other)

## Top Skills

SAS  
Statistical Modeling  
Data Mining

## Languages

Spanish (Elementary)  
French (Elementary)  
English (Native or Bilingual)

## Certifications

Introduction to R  
Teach English Now! Foundational Principles  
Intermediate R Course  
Data Scientist with R Track

## Publications

Geographic variation within the United States in the incidence of hepatocellular carcinoma.  
A practical screening tool for anxiety and depression in patients with chronic breathing disorders.  
Method to develop health care peer groups for quality and financial comparisons across hospitals.

# Harlan Nelson

Posit Fan  
Indianapolis, Indiana, United States

## Summary

I can solve analysis problems, make systems more efficient and simple, and explain and train. I can find new ways to do things and introduce new current methods to your processes.

I am not as deadline focused as many would like and have a tendency to diverge off scope at times when implementing solutions. Relaxing some business requirements is necessary to give the room to make changes. Reworking solutions is the process of learning. However, keeping a linear path to a fixed deadline is not my skill.

I have significant experience using SAS, R, Python, Pig, SQL and SPSS Data Modeler. I am excited about Bayesian methods and really like R Shiny. I am interested in modern analysis methods in data mining, machine learning and statistical analysis. I am available to help solve any programming or analysis problems you might have, or for education.

---

## Experience

Indiana University Health  
Senior Data Scientist Analyst  
February 2021 - Present (3 years 11 months)  
Indianapolis

Lucid Services Group  
Data Analyst  
July 2019 - Present (5 years 6 months)  
Indianapolis, Indiana, United States

Develop patient cohorts for healthcare research.  
Help implement an automated interface for cohort development  
Work with students completing capstone projects.  
Utilize Azure Data Warehouse and R for analysis and cohort building.

## Sollers College

### Instructor

December 2017 - July 2019 (1 year 8 months)

Menlow Park, NJ

I have taught for their Data Science program. I have taught some sessions in Data Mining with R and for their Background Module.

## USA English Audiovisual School

### Instructor

May 2014 - July 2019 (5 years 3 months)

400 Park Ave. 2ND Floor Plainfield, NJ 07060

I teach English Classes as needed. I currently only substitute as needed when a teacher is out.

## Kelly Services

### Real World Data Analyst

June 2018 - October 2018 (5 months)

- Created cohort extractions and summaries from 10 Terra-bytes of Truven data using SAS SPDE and R.
- Created a production version of a suite of SAS code for cohort extraction and patient attrition.
- Converted specifications such as DX, procedure and drug names from being hard coded in SAS to being read from an Excel Spreadsheet. The cohort extraction could then be specified by listing the DX codes, drug names and procedure names in an Excel Spreadsheet instead of modifying SAS code.
- Instead of running QC by peppering the SAS code with print statements, I wrote an R notebook that read in the data produced by SAS and produced QC tables.
- Implemented SAS code management using BitBucket. This allowed recovery of past code versions.
- Wrote a parser in R to search the SAS log for problem keywords as indicated in a separate text file.
- Wrote a parser in R to extract all table creation notes from the SAS log into an organized table showing dataset progression.
- Implemented YAML for the documentation format in the SAS code headers.
- Wrote a YAML parser in R to compile SAS documentation headers to a separate report. This allowed automatic extraction of all SAS program documentation headers into a consolidated project report. This is useful for QC.

- Wrote interactive data exploration using R Shiny to dynamically view cohort summaries.
- Wrote project summaries using R Bookdown.

## Oxford International

### SAS Programmer/Statistician

July 2017 - September 2017 (3 months)

Warren, NJ

- Converted market size calculation of anti-emetic drugs from Excel and SAS to just SAS.
- Conducted analysis to verify, from Empirical data, that prescription doses were consistent with recommended doses.
- Simplified and standardized disorganized calculation methodology.
- Identified steps in a calculation process that did not access validated data and recommended running all calculations against a validated central database.

## Agile Outcomes Research, Inc

### SAS Programmer

September 2016 - January 2017 (5 months)

Work from Home

Wrote SAS code for the analysis of a COPD study comparing LABA and SABA use in a nursing home.

## INSYS Group Inc

### Data Analyst

August 2014 - July 2016 (2 years)

New Jersey

- Worked on site for a telecommunications client mining data in a Hadoop environment. Used Pig, Python, SPSS Data Miner, R and HiveQL.
- Associated mobile phone data device characteristics with propensity to service calls.
- Compiled battery drain data and looked for correlations related to device characteristics such as firmware versions.
- Identified devices with bad batteries.
- Looked for device factors related to device return.
- Applied Gamma-Poisson observed-over-expect methods to mobile phone data.

## inVentiv Health Clinical

## Principal Statistician

August 2013 - May 2014 (10 months)

- Supported client Pfizer.
- Created reports for health outcomes research using SAS and Excel
- Pulled extracts from the Humedica electronic medical records and the KIMS registry to conduct health outcomes research.
- Studied ulcerative colitis and adult human growth hormone deficiency.
- Wrote a programmer request and helped write an SAP.

## MaxisIT Inc.

### Contractor

November 2012 - May 2013 (7 months)

New Jersey

- Supported real world analytics in the CEHO (Cost Effectiveness and Health Outcomes) department of client Sanofi-Aventis.
- Utilized the health care claims data called "The MarketScan® Commercial Claims and Encounters and Medicare Supplemental Databases from Truven Health Analytics", "Optuminsight IMPACT multi-payer Database" and "INGENIX LABRX."
- Proposed and tested a modification to in-house data structure that would allow cohort identification in minutes, saving hours of computing time.
- Proposed and tested the use of SAS asynchronous processing.
- Helped write some SOPs.

## CoreIT

### Statistical Programmer

June 2012 - October 2012 (5 months)

Produced reports from Medicare part A and B Claims Data using SAS and DB2 as directed by the client: Blue Cross BlueShield of South Carolina. (Palmetto GBA).

## Department of Veterans Affairs

### Statistician

January 2012 - April 2012 (4 months)

Indianapolis, Indiana Area

- Conducted Health Services Research for the Department of Veterans Affairs using primary data and Electronic Health Records.
- Worked in a "Center of Excellence in Evidence Based Practice (CIEBP)." Measured clinical compliance to standards of care for Stoke, and Transient ischemic attacks (mini-strokes).

## SDI

### Statistician

May 2010 - December 2011 (1 year 8 months)

Plymouth Meeting, PA

Worked at SDI which was bought by IMS in 2011. Was converted from contract to a permanent position. Extracted and compiled health care claims data and Electronic Medical Records using Oracle PL/SQL and SAS. Helped develop analytical methods that used drug regimens to identify cancer stages. Multinomial and Discrete Choice models were used. Learned and used the SDI method (designed by Paul Smith) of projecting nation wide health measures using indirect estimators. Conducted market campaign analysis. Conducted data mining using regression trees and random forests with the "rattle package in R. Wrote HOWTO notes explaining use of R and WPS to access Oracle Data. Also gave a talk on using R with Oracle.

## inVentiv Clinical Solutions LLC

### Statistician

March 2008 - November 2009 (1 year 9 months)

- Helped harmonize methodology for, program, and implement a validated system in SAS Drug Development (SDD) used to estimate the number of patients exposed to marketed molecules. These estimates were used in regulatory reports submitted to the FDA. Wrote an analysis plan and collaborated with other contractors to implement these requirements in the SAS Drug Development platform (SDD).
- Created data sets and reports for clinical trials.

## Volt Life Sciences

### Project Statistician

March 2006 - March 2008 (2 years 1 month)

- Helped harmonize methodology for, program, and implement a validated system in SAS Drug Development (SDD) used to estimate the number of patients exposed to marketed molecules. These estimates were used in regulatory reports submitted to the FDA. Wrote an analysis plan and collaborated with other contractors to implement these requirements in the SAS Drug Development platform (SDD).
- Researched and presented recommendations of statistical methodology for automated signal detection. The Proportional Reporting Ratio (PRR) and empirical Bayesian methods were reviewed using an in-house Safety Database and the FDA's Spontaneous Adverse Event Reporting System (AERS) Database.

- Collaborated with other scientists to develop new analytical capabilities: wrote simulation programs in SAS using Monte Carlo methods to compare performance characteristics of statistical methods used in meta-analysis.
- Wrote simulation programs in SAS to calculate sample size needs for a matched case control clinical trial design.

Laura A. Petersen, M.D., M.P.H., F.A.C.P.

Contractor

October 2005 - January 2006 (4 months)

I helped finish up a project for the current director of the Center.

- Contributed statistical modeling support for health outcomes research. Focused on quality and cost of managed health care.
- Provided statistical programming support for team led research. This team provided budget recommendations and quality monitoring to VA hospital networks using analysis of VA health care databases.
- Developed statistical models used as tools for health care quality monitoring.

Department of Veterans Affairs

Statistical Analyst

August 1998 - September 2005 (7 years 2 months)

- Consulted with medical researchers and other team members to provide SAS programming support, statistical modeling support, analysis data set development, and database support.
- Wrote SAS data base queries against the VA electronic medical records.
- Developed innovative statistical methodology for health outcomes research.
- Implemented DxCG and ACG based healthcare cost and quality statistical models.
- Implemented cluster methodology to compare hospital facilities on cost and quality.
- Used multivariate and mixed models to improve efficiency of health care with provider profiling.
- Helped develop methodology for health payment systems at the Veteran's Health Administration through work with disease burden adjustment models. The electronic medical records at the Veteran's Health Administration (VHA) are a single payer version of Medical Claims records. They are used to determine reimbursement, but the VHA uses a disease burden model for payment. It is scientifically more advanced than the fee for service associated with claims records. These records are more complete than claims records and the payment system used by the VHA allows it to spend most of its research dollars improving care.

## Northern Illinois University

### Graduate Assistant

August 1995 - May 1998 (2 years 10 months)

DeKalb, IL

- Assisted in the courses Statistics, Finite Math, and Calculus.
- Conducted three recitations per week, Graded Homework.
- Worked in the Statistics Consulting Lab

## Trinity International University

### Lecturer

October 1993 - October 1995 (2 years 1 month)

Deerfield, IL

- Taught, with full classroom responsibility, Intermediate Algebra, Statistics, Pre-Calculus, and Calculus I.
- Also helped write a manual used to integrate SPSS into an introductory statistics class.

---

## Education

### De Paul University

Master's Degree, Mathematics · (1990 - 1992)

### Northern Illinois University

None, Statistics · (1995 - 1998)

### Trinity International University

Bachelor's degree, Math and Computer Programming · (1985 - 1989)