

# Ben Barnard

Data Scientist



(817) 937 - 6705



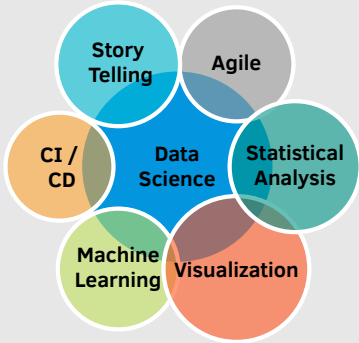
benbarnard87@gmail.com



BenBarnard

## Skills

### Overview



### Programming

C++ • SQL • Python • SAS  
shiny • R • rmarkdown  
ggplot2 •  $\LaTeX$

## Education

2018	<b>Ph.D., Statistical Sciences</b> Dissertation: A Power Contrast of Tests for Homogeneity of Covariance Matrices in a High-Dimensional Setting	Baylor University
2016	<b>M.S., Statistical Sciences</b>	Baylor University
2013	<b>M.S.Ed., Exercise Physiology</b>	Baylor University
2011	<b>B.S., Exercise Science</b>	University of Texas at Arlington

## Experience

May 2018 - Present	<b>Data Scientist</b> <ul style="list-style-type: none"><li>Model development lead for Human Capital Key Risk Indicators using time series models.</li><li>Developer of 20 internal R packages using agile development with 2-week sprints and daily asynchronous stand-ups in Skype chat groups.</li><li>Enterprise Product Steward for R and Rtools monitoring intake and maintenance of new versions, and mitigating security vulnerabilities for ~2500 installs across the enterprise.</li><li>Lead transformation of team to use software development life-cycle appropriate tools such as GitHub, Jenkins, and Artifactory.</li><li>Model developer for turnover model using random forest to inform a Bayesian joint estimable logistic regression and cox proportional hazard model in stan.</li></ul>	Wells Fargo
Aug 2017 - Present	<b>Statistics Instructor</b> <ul style="list-style-type: none"><li>Coordinator for design and instruction of Applied Statistics for Health Sciences courses.</li><li>Quality matters and Team Based Learning instruction certifications.</li><li>Developer of tumor image classification model using a convolution neural network in tensorflow.</li></ul>	University of South Alabama
Mar 2017 - Aug 2017	<b>Biostatistics Intern</b> <ul style="list-style-type: none"><li>Model developer of physician quality incentive programs random forest model.</li><li>Statistical analysis of Necrotizing Enterocolitis retrospective study using logistic regression mixed model in PROC GLIMMIX.</li><li>Bayesian sample size simulation study in stan for Family, Food and Fun program.</li></ul>	Baylor Scott & White Health
May 2013 - Aug 2017	<b>Graduate Assistant</b> <ul style="list-style-type: none"><li>Developer for conversion rate key performance indicators on graduate applications.</li><li>Model developer for student success gradient boosting machine model using XGBoost.</li><li>Designer of A/B testing for the Graduate School website and application pages.</li><li>Model developer for applicant and graduate student segmentation clustering models using k-means clustering.</li><li>Developer for automated reporting on applications and enrollment data using D3.js, SAS reporting studio, Shiny server.</li><li>Model developer for athlete injury prediction model using recurrent neural network in tensorflow.</li><li>Statistical analysis on paleosol composition using semi-parametric models in PROC TRANSREG.</li></ul>	Baylor University

## Presentations

- Feb 2019      **Conference on Statistical Practice**  
Automated Building and Storing Frozen Data in R Packages Using Travis and Drat.
- Mar 2019      **University of Miami Biostatistics Department**  
Automated Building of R Packages Using Travis and Drat for data, shiny, and websites.
- Aug 2019      **Joint Statistical Meetings**  
A Journey Teaching Applied Statistics for Health Sciences in an Asynchronous Team Based Learning Format using Data Science Ideas.
- Jan 2020      **rstudio::conf** (Accepted)  
The good, the bad and the ugly: What I learned while consulting across the business as a data scientist.

## R Packages

- Apr 2017 - Present      **rWishart**  
An expansion of R's 'stats' random wishart matrix generation. This package allows the user to generate singular, Uhlig and Harald (1994), and pseudo wishart, Diaz-Garcia, et al.(1997), matrices. In addition the user can generate wishart matrices with fractional degrees of freedom, Adhikari (2008), commonly used in volatility modeling. Users can also use this package to create random covariance matrices.
- Jan 2017 - Present      **mvmonitoring**  
Use multi-state splitting to apply Adaptive-Dynamic PCA (ADPCA) to data generated from a continuous-time multivariate industrial or natural process. Employ PCA-based dimension reduction to extract linear combinations of relevant features, reducing computational burdens. For a description of ADPCA, see the 2016 paper from Kazor et al. The multi-state application of ADPCA is from a manuscript under current revision entitled "Multi-State Multivariate Statistical Process Control" by Odom, Newhart, Cath, and Hering, and is expected to appear in Q1 of 2018.
- Jun 2016 - Present      **covTestR**  
Testing functions for Covariance Matrices. These tests include high-dimension homogeneity of covariance matrix testing described by Schott (2007) and high-dimensional one-sample tests of covariance matrix structure described by Fisher, et al. (2010). Covariance matrix tests use C++ to speed performance and allow larger data sets.
- Jul 2016 - Present      **likelihoodExplore**  
Provides likelihood functions as defined by Fisher (1922) and a function that creates likelihood functions from density functions. The function are meant to aid in education of likelihood based methods.

## Service

- Jan 2019 - Dec 2021      **Council of Sections** American Statistical Association  
Executive council member for the Section of Statistical Programmers and Analysts representing the section on the council of sections.
- Oct 2018 - Present      **Peer Journal Reviewer** Journal of Statistics Education  
Reviewer in the areas of Data Science, multivariate, and health sciences statistics education.