Homework 1
APPM 5650 Fall 2021
Randomized Algorithms

Due date: Friday, Aug 27 2021, in class
Theme: Background

Instructor: Prof. Becker

Instructions Collaboration with your fellow students is allowed and in fact recommended, although direct copying is not allowed. Please write down the names of the students that you worked with. The internet is allowed for basic tasks only, not for directly looking for solutions.

An arbitrary subset of these questions will be graded.

Problem 1: Skim the first 2 sections of Michael Mahoney's "Randomized algorithms for matrices and data" (2011, part of the Foundations and Trends® in Machine Learning series). DOI link and arXiv link. No response necessary.

Problem 2: Read chapter 1 "Introduction" in "Randomized Numerical Linear Algebra: Foundations & Algorithms" (P.G. Martinsson and J. Tropp, 2020, Acta Numerica vol 29). DOI link and arXiv link. No response necessary.

Problem 3: Fill out the Google Forms survey (link in Canvas)

Problem 4: Read the "Appetizer" chapter of Roman Vershynin's 2018 "High-Dimensional Probability: An Introduction with Applications in Data Science". Book website and a free PDF on Vershynin's website.

In the last equation of the proof of Theorem 0.0.2 (Approximate Caratheodory's theorem), it is shown that for a random variable Y (where $Y = x - \frac{1}{k}\sum_{j=1}^{k} Z_j$ in Vershynin's notation), it holds that $\mathbb{E}\,\|Y\|_2^2 \leq \frac{1}{k}$, and he therefore concludes that there is a realization of $Y = y$ that satisfies $\|y\|_2^2 \leq \frac{1}{k}$.

Please turn in a brief response that justifies Vershynin's conclusion. This is a case of Erdös's Probabilistic Method, and you may read online about it, but your response should be in your own words.

Note: you only need to prove this for the case of **discrete random variables**. The case of *continuous* or even *absolutely continuous* random variables is less clear.