

# An Analysis of Exercising Behavior in Online Populations

David Jurgens, James McCorrison, Derek Ruths



jurgens@cs.mcgill.ca, james.mccorrison@mail.mcgill.ca, derek.ruths@mcgill.ca

## Introduction and Motivation

Exercise is an important component of health. Previous studies of exercise and health have been stymied by the lack of accurate records of exercising activities. In our study, we overcome this limitation by using data from Fitocracy, a gamified workout tracking site that precisely records users' workout histories using over 1,000 unique exercises and where users self-report their age, gender, and height. We analyze nearly half a million users' histories to identify exercising behavior trends and how those behaviors vary by age and gender.

## Data

All of a user's activities, profile, social data, and group memberships on Fitocracy were crawled over a six-month period to acquire the complete profiles and workout histories of 441,034 users. Ultimately, 188,265 users recorded at least one workout, with the total dataset comprising 3,130,276 workouts (14.3M activities) over nearly a four year span from February 2011 to January 2015.

## How does exercising behavior vary with age and gender?

**Methodology:** The activities performed by different age groups provide a insightful view into how exercise behavior changes over time. We divided users by gender and into seven age ranges. Within each gender-age cohort, we computed the probability that a user in that cohort records each exercise and then sorted all exercises according to their average probability of being performed. Below, we show the ten exercises that are most likely to be performed by an individual from each cohort.

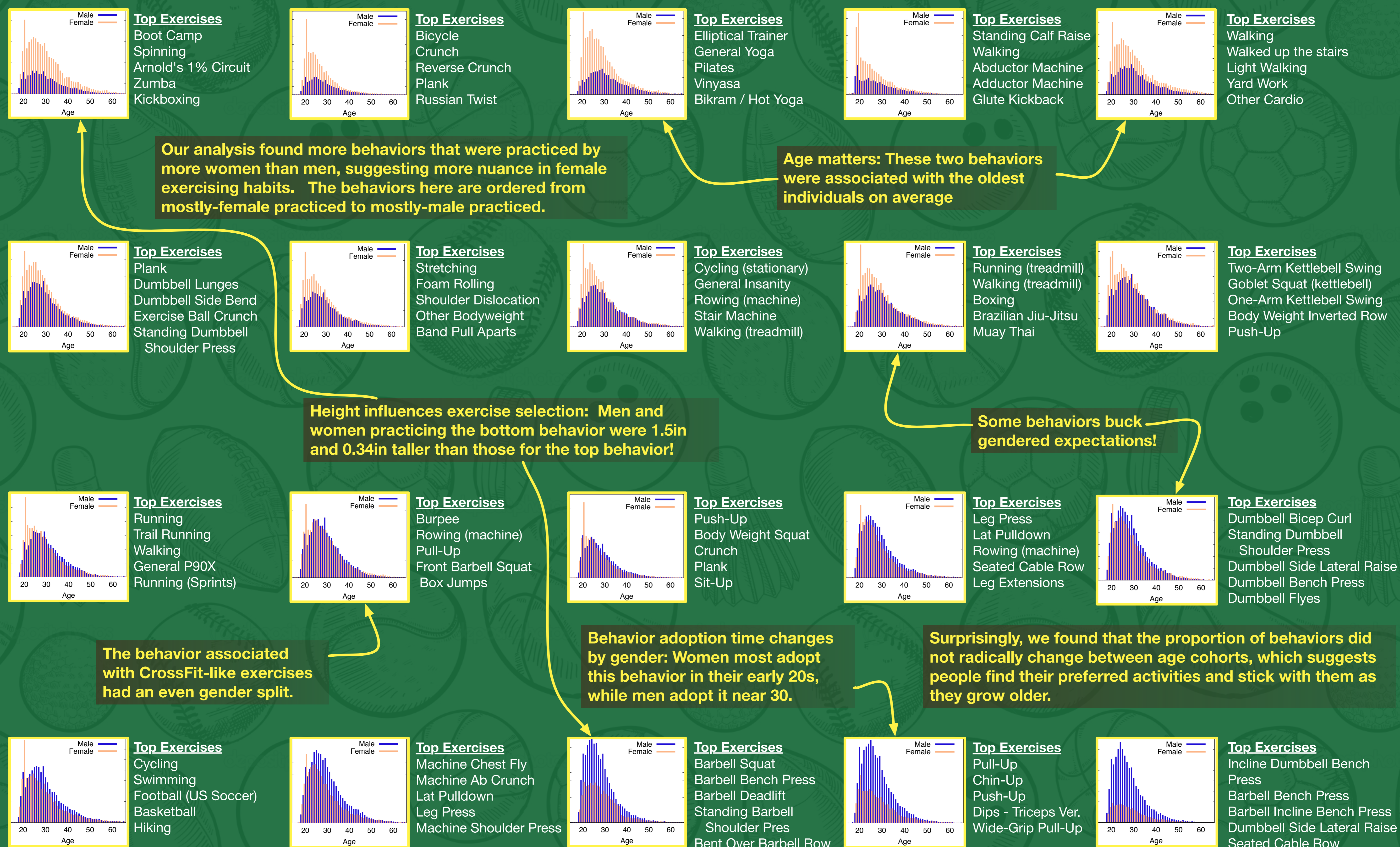
	Ages 0-20	Ages 20-25	Ages 25-30	Ages 30-35	Ages 35-40	Ages 40-50	Ages 50+
<b>Women</b>	Walking	Walking	Walking	Walking	Walking	Walking	Walking
	Running	Running	Running	Running	Running	Running	Walking (treadmill)
	Crunch	Elliptical Trainer	Elliptical Trainer	Elliptical Trainer	Elliptical Trainer	Elliptical Trainer	Running
	Plank	Crunch	Running (treadmill)	Running (treadmill)	Running (treadmill)	Walking (treadmill)	Elliptical Trainer
	Body Weight Squat	Running (treadmill)	Push-Up	Push-Up	Walking (treadmill)	Running (treadmill)	Cycling (stationary)
	Push-Up	Push-Up	Plank	Plank	Push-Up	Push-Up	Walked up the stairs
	Elliptical Trainer	Plank	Crunch	Crunch	Body Weight Squat	Plank	Cycling
	Running (treadmill)	Body Weight Squat	Body Weight Squat	Body Weight Squat	Plank	Cycling (stationary)	Swimming
	Jumping Jacks	Cycling (stationary)	Cycling (stationary)	Walking (treadmill)	Crunch	Crunch	Running (treadmill)
	Sit-Up	Stretching	Stretching	Cycling	Cycling (stationary)	Stretching	Push-Up
<b>Men</b>	Push-Up	Push-Up	Running	Running	Running	Running	Walking
	Barbell Squat	Running	Push-Up	Push-Up	Walking	Walking	Running
	Barbell Bench Press	Barbell Squat	Barbell Squat	Walking	Push-Up	Push-Up	Cycling
	Running	Barbell Bench Press	Barbell Bench Press	Barbell Squat	Running (treadmill)	Cycling	Push-Up
	Barbell Deadlift	Barbell Deadlift	Walking	Barbell Bench Press	Cycling	Running (treadmill)	Elliptical Trainer
	Pull-Up	Pull-Up	Barbell Deadlift	Running (treadmill)	Barbell Bench Press	Elliptical Trainer	Cycling (stationary)
	Dumbbell Bicep Curl	Dumbbell Bicep Curl	Pull-Up	Cycling	Elliptical Trainer	Cycling (stationary)	Running (treadmill)
	Walking	Walking	Running (treadmill)	Elliptical Trainer	Barbell Squat	Barbell Bench Press	Walking (treadmill)
	Crunch	Crunch	Cycling	Barbell Deadlift	Cycling (stationary)	Barbell Squat	Light Walking (secondary)
	Overhead Press	Sit-Up	Dumbbell Bicep Curl	Pull-Up	Pull-Up	Crunch	Walked up the stairs

**Women's exercise selection focuses on bodyweight movements and cardio and is remains largely similar across age cohorts.**

**Men's exercise selection initially focuses on strength training but then converges with women's exercise choices.**

## What exercising behaviors are practiced by subpopulations?

**Methodology:** To identify the underlying behaviors from people's activities, we train a Latent Dirichlet Allocation (LDA) model on users' exercising data. We model each individual's history as a reflection of that person engaging in just a few behaviors, where a behavior selects for certain exercises with higher frequency. Below, we show the demographics of the population engaging in the behaviors identified from a 20-behavior LDA-model.



Check out our interactive behavior demo at <http://networkdynamics.org/resources/exercise>