---

layout: default

title: Research Interests

---

<div class="blog">

<h1>Data Analytics Blog #1: What's going on with Michigan Football?</h1>

<img src="{{ site.baseurl }}/static/michigan\_football.png" width="800" align="center">

Jim Harbaugh was supposed to be the savior of Ann Arbor. The man who would finally bring Michigan back to its glory days, its rightful place on top of the Big Ten, and a National Championship. But these days, many people are having a hard time imagining Michigan even winning its division, much less the Big Ten or a National Championship.

So what in the world is going on with the Michigan Wolverines?

I'm a researcher and data analyst, and so I thought we might be able to derive some insights into the Michigan Wolverines' struggles using data analytics. If you're interested in the data or code I used to derive these insights, you can go to my GitHub page <[link]> to access all the data and code.

<h2>Is having an elite offense better than having an elite defense?</h2>

First, defining the problem:

Michigan is actually quite good against decent opponents. They're 7-3 and have outscored teams ranked 10-25 by an average score of 31-14. They've had a top-10 recruiting class three of the past four years and a top-10 defense four years in a row. But, their elite defense and elite talent have not been nearly good enough when it comes to playing other elite teams. Instead, Michigan is 1-9 against top-10 opponents, has been outscored 34 to 19 on average, and has only scored 30+ points once.

Clearly when you're being outscored 34 to 19 there are issues that need to be fixed on both offense and defense, but I am going to test the hypothesis that, contrary to the old adage that "Defense wins championships", an elite offense is crucial to winning games against elite opponents. In other words, if you can only have an elite offense or an elite defense,

which gives you a better shot at beating other elite teams?

To do this, I am using data from the CollegeFootballData.com API which contains S&P+ team ratings as well as game outcomes for each team.

It turns out that if you have a top-10 offense, you win about 47% of the time agains a top-10 ranked opponent. If you have a top-10 defense, you only win about 40% of time. Similarly, if you have an elite offense but not an elite defense, you win about 40% of the time, and if you have an elite defense but not an elite offense, you win about 33% of the time.

<[insert Figure1.png here]>

So that suggests that having an elite offense is slightly more advantageous than having an elite defense. But Jim Harbaugh hasn't even won 33% of the time against elite opponents, so clearly there's more going on than just lacking an elite offense.

Some of that is luck (especially the 2015 Michigan State and 2016 Ohio State games-- winning those two games would get Michigan to a 33% win rate), but some of that might also be the antiquated offensive philosophy that Harbaugh has emphasized the past few years. Having big TEs and mauling offensive linemen helps you against teams that you can out-muscle, but seems less effective against teams that also have big, athletic, and talented players. Jim Harbaugh seems to have conceded this point, hiring Josh Gattis in the offseason to modernize the offense and convert it from a pro-style offense to a spread-offense. Gattis' motto in the offseason was #speedinspace and the goal was to install an offense that would spread the defense out and

give the ball to Michigan's athletes in space to allow them to make big plays.

But how well has this worked out so far? In the wake of mediocre offensive performances against Army and Wisconsin, people have begun to question "whether the offense lacks identity", "if the system has been fully implemented yet", or "if the system is too complicated for college players".

<h2>Growing pains when implementing a new offense</h2>

To answer these questions, I'm going to be scraping play-by-play data from MGoBlog.

This is a great dataset because Brian Cook, the founder of MGoBlog, has painstakingly coded each play for offensive and defensive formation, play call, and number of RBs, WRs, and TEs on the field.

Has Michigan truly implemented a spread offense?

The first way I looked at this question is to ask, what percentage of the time is Michigan putting its players in space? Lots of situations and play calls put players in space, but I chose to focus on two fundamental aspects of spread offenses:

1. Percentage of non-goal-line plays with 3+ WRs

2. Percentage of non-goal-line run plays that utilize a read option, QB run, or run-pass option

Michigan utilizes a 3-WR or 4-WR set on 65% of non-goal-line plays, including 72% of non-goal-line pass plays and 55% of non-goal-line run plays. This is pretty good, but still doesn't represent a full conversion to a spread offense, as they are utilizing 2 TEs on the field 35% of the time. So this suggests that they have “mostly” implemented a spread offense.

<[insert Figure2.png here]>

Their implementation of spread concepts such as the read option, QB run, or run-pass option, on the other hand, is lagging far behind. Some of this might be due to Shea Patterson and Dylan McCaffrey being injured, but my count Michigan has only run a play that utilized a read option, run-pass option, or designed QB run on 13% of plays.

<[insert Figure3.png here]>

This points to a failure to install pieces of the offense that are fundamental to how spread offenses work. It lends some credence to the idea that Michigan’s offense lacks an identity, as they have spread personnel on the field but aren’t running spread plays.

In the last game against Rutgers we saw Michigan implement the waggle more frequently (23% of plays), which was very successful at getting the QB into space and making a defender cover the QB (opening up other players). This suggests that in earlier games the spread offense might not have been fully operational. In the future, it would be nice to see more read option or run-pass option plays mixed in with the waggle to make the defense truly respect the threat of a QB run.

<h2>How well is Michigan’s new offense working?</h2>

Beyond simply asking whether Michigan is actually implementing a true spread offense, we can also ask which features of their offense (spread or not spread) are working the best. To assess this, I looked at two statistics. The first is average yards per play (YPP), which tells you on average how many yards the offense gains on each play type. The second is success rate, which is a little more complicated but basically assigns a 1 if the offense gained a "good" amount of yards and a 0 if the offense didn't gain a "good" amount of yards. "Good" in this context depends on the situation. If the play is a first-down play, a "good" amount of yards is gaining at least half the yards needed for a first-down (e.g., 6 yards on 1st and 10). If the play is a second down, a "good" amount of yards is gaining at least 70% of the yards needed for a first-down (e.g., 5 yards on 2nd and 7). If the play is third down, a "good" amount of yards is gaining at least 100% of the yards needed for a first down (for obvious reasons).

In college football, anything over 5 yards/play is considered a good run play average, and anything over 8 yards/play is considered a good pass play average. For success rate, a team is doing well if they get a "good" amount of yards on more than 45% of plays (run or pass).

[NEED TO ADD TEXT DESCRIBING YPP AND SUCCESS RATE FOR SPREAD-TYPE PLAYS]

Focusing on plays where Michigan had 3 or more WRs on the field (which were 65% of plays), Michigan was much more effective on both run and pass plays where there were 3 or more WRs on the field. This is due to 2 reasons. First, having 3 WRs on the field more effectively spread the defense out, opening more holes for the RBs.

<[insert Figure4.png here]> yards/play and success rate for run/pass with 2 WR or 3+ WR

Second, Michigan's WRs are incredibly good. In particular, Nico Collins is averaging 12.9 yards/play, catching 67% of the balls thrown to him, with a very impressive 63% success rate.

For context, in 2017 (the most recent year I have stats for individual players), only 3 WRs (who had at least 5 targets/game) in the entire country averaged more than 12 yards/play with at least a 60% success rate. Perhaps, more impressively, Michigan's 2nd, 3rd, and 4th best WRs (Black, Bell, and Peoples-Jones) average at least 8 yards/play with a success rate over 50%.

For context, Oklahoma State was the only team in the country in 2017 with 4 WRs who averaged at least 8 yards/play with a success rate over 50%, and they had the 5th best offense in the country that year. So it makes sense that the team would be better when they have at least 3 WRs on the field.

<[insert Figure5.png here]> yards/play and success rate for top-4

But here's the thing: Only 57% of passes are going to those four players. Diversity is a good thing, and a good offense will need to incorporate passes to TEs and RBs, but if you have four of the best WRs in the country, you need to get the ball to them more than 19 times/game. The good news is that 73% of passes in the last game against Rutgers went to Collins, Black, Bell, or Peoples-Jones.

<[insert Figure6.png here]> percentage of passes thrown to top-4 in each game

<h2>Recommendations for future success</h2>

So for the second time, we see some evidence that Michigan is already starting to implement the thing that my analyses suggest will make the offense better (making the QB a run threat and throwing to your top WRs). Tomorrow will be an important test to see whether Michigan can stick to this plan and have success against a better defense, but I am confident that Michigan's offense will be better if they make the QB a legitimate run threat,

utilize more 3 WR sets, and throw the ball to their WRs in space.

In my next blog post, I will revisit these recommendations and analyze how well these three features were implemented in Michigan's gameplan, how successful they were, and also test some of the other questions about Michigan’s offense, including whether it’s too complicated for college players and how often you need a QB run to keep the defense honest.