Sports Reference Blog

What the Heck is Corsi? A Primer on Advanced Hockey Statistics

Posted by Jonah Gardner on October 13, 2016

Good news for fans of zambonis, fighting, and the greatest video game of the 1990s: the NHL has finally returned! After a wild season last year, there are all kinds of juicy storylines to follow this year. Can the Pittsburgh Penguins become the first back-to-back Stanley Cup winners since the Detroit Red Wings of the 1990s? How will the San Jose Sharks bounce back from coming so close and falling short. Will Alex Ovechkin reach 1,000 goals? Can Connor McDavid build upon a promising rookie year and live up to the hype? What round of the Eastern Conference Playoffs will the Washington Capitals be eliminated in this year (I kid, I kid)?

This blog post will seek to answer none of those. Instead, this week, I wanted to dig into one of the major trends that's been sweeping across the NHL the last few years, among fans and front offices alike. I'm talking, of course, about the rise of advanced statistics.

If you're a sports fan, you're probably at least vaguely familiar with *Moneyball* and the advanced stat wars in baseball. And you may have read articles about how thinkers in other sports, like basketball, have used similar principles to deepen their understanding of the game. This movement has reached hockey in recent years, as researchers have uncovered several new ways of understanding the game beyond the traditional stats like goals, assists, and plus/minus. These new analytics can help us understand why a team is over or under-performing, and whether that performance is sustainable. They can also help us appreciate unsung players who do more for their team than we may realize, because they don't put up flashy traditional numbers.

So, with that in mind, here's some of the basics to get you started in the world of advanced hockey stats.

Corsi and Fenwick

Corsi and Fenwick are the closest thing you'll find to crossover stats in the advanced hockey analytics world. Tim Barnes, one of the people who helped develop, popularize, and spread those stats, works for the Caps, while everywhere from Vice to (R.I.P.) Grantland has written about them. But, despite their fancy names, both stats are pretty simple.

At their heart, Corsi and Fenwick are a proxy for measuring possession, since actual possession time is not recorded by official statisticians. The idea behind them is quite intuitive: the more you possess the puck, the more shots you'll be able to create and the more likely you'll be to score.

Starting with Corsi, there are two kinds of Events: a Corsi For Event and a Corsi Against Event. Again, despite the names, those are quite simple: a Corsi For Event is a shot attempt by your team, a Corsi Against Event is a shot attempt by your opponent. Corsi Events include all shot attempts, regardless of whether they're saved, blocked, off-target, or scored.

However, the raw Corsi For and Corsi Against numbers don't tell us too much. Instead, what we want to know is whether a team is attempting more shots than they're allowing. To do that, we can calculate a team's Corsi For Percentage by dividing their number of Corsi For Events by the team's total number of Corsi Events, both For and Against. If you're over 50%, things are probably going pretty well. Here, for instance, are the Top 10 teams in CF% at even-strength in 2015-16:

Query Results Table

			Corsi (EV)
Rk	**Tm**	**Season**	**CF%**
1	LAK	2015-16	56.1
2	PIT	2015-16	52.8
3	ANA	2015-16	52.4
4	NSH	2015-16	52.3
5	DAL	2015-16	52.3
6	TBL	2015-16	51.8
7	STL	2015-16	51.8
8	DET	2015-16	51.8
9	TOR	2015-16	51.6
10	WPG	2015-16	51.5

Provided by Hockey-Reference.com: View Original Table
Generated 10/12/2016.

Hey, look; there's the champs at number two! In fact, all eight of the top eight Corsi For Percentage teams made the playoffs last year. And while the Sharks, who won the West, and the Caps, who had the best regular season record, didn't make the Top 10, both teams finished over 50%.

What about Fenwick? It actually measures the same thing, with one slight distinction. Fenwick doesn't count blocked shots as Fenwick Events. If you think blocking and avoiding blocked shots are skills that should be separated out from the general possession calculation, then Fenwick is the stat for you. Again, the Fenwick For Percentage leaderboard includes a lot of teams that you may remember watching last May and June:

Query Results Table

			Fenwick (All)
Rk	**Tm**	**Season**	**FF%**
1	LAK	2015-16	54.8
2	NSH	2015-16	53.3
3	SJS	2015-16	53.0
4	ANA	2015-16	52.7
5	PIT	2015-16	52.5
6	DAL	2015-16	52.2
7	CAR	2015-16	51.7
8	TOR	2015-16	51.2
9	WSH	2015-16	51.2
10	PHI	2015-16	51.1

Provided by Hockey-Reference.com: View Original Table
Generated 10/12/2016.

That's how these stats work on a team level, but you can also apply them to individual players. If you see a Corsi For or Fenwick For Percentage for an individual skater, that's measuring the ratio of shot attempts by his team while he was on the ice. It's like plus/minus with shots instead of goals. Here were the Top 10 players in CF% last year (with a minimum of 40 games played):

Query Results Table

				Corsi (EV)
Rk	**Player**	**Pos**	**Tm**	**CF%**
1	**Nick Shore**	C	LAK	61.4
2	**Tyler Toffoli**	C	LAK	59.3
3	**Milan Lucic**	LW	LAK	59.0
4	**Pavel Datsyuk**	C	DET	58.3
5	**Drew Doughty**	D	LAK	58.1
6	**Brayden McNabb**	D	LAK	58.0
7	**Mathieu Perreault**	C	WPG	57.9
8	**Mike Ribeiro**	C	NSH	57.9
9	**Dustin Brown**	RW	LAK	57.8
10	**Tomas Tatar**	LW	DET	57.7

Provided by Hockey-Reference.com: View Original Table
Generated 10/12/2016.

That's a lot of Los Angeles Kings! Indeed, looking at that list brings up an obvious question with this framework: If a player plays for an excellent Corsi team, won't his own Corsi Number be inflated, regardless of his indivdual performance? And that's true, but there's another stat to help control for that.

Instead of using regular Corsi or Fenwick For Percentage, I like to look at Relative Corsi/Fenwick For Percentage instead. These relative percentages measure the change in Corsi or Fenwick when a player is on the ice versus when he's on the bench. If you're a bad player who is coasting off your teammates or a good player being dragged down by the skaters around you, the relative stats will do a better job of discovering that:

Query Results Table

				Corsi (EV)
Rk	**Player**	**Pos**	**Tm**	**CF% rel**
1	**Mathieu Perreault**	C	WPG	8.0
2	**Patrice Bergeron**	C	BOS	7.9
3	**James van Riemsdyk**	LW	TOR	7.7
4	**Mike Ribeiro**	C	NSH	7.2
5	**Nino Niederreiter**	RW	MIN	7.2
6	**Brad Marchand**	LW	BOS	6.9
7	**Eric Staal**	C	TOT	6.7
8	**Erik Karlsson**	D	OTT	6.5
9	**Nick Shore**	C	LAK	6.4

				Corsi (EV)
Rk	Player	Pos	Tm	CF% rel
10	Evander Kane	LW	BUF	6.3

Provided by Hockey-Reference.com: View Original Table
Generated 10/12/2016.

So, the Relative Corsi For Percentage leaderboard includes players like Patrice Bergeron and Erik Karlsson, who were missing from the regular CF% leaderboard because they played for mediocre teams last year.

These numbers represent an approximation of possession time, and possession time means more chance to score, and less chance for your opponent to score. If you're "driving possession," it's a good thing. Of course, you say, it's better to actually score and prevent goals, as opposed to shots. And that's true. But because of the small sample size of goals for/against, it's not always a fully reliable indicator.

PDO

If Corsi and Fenwick represent a general way of evaluating how a team is actually performing underneath the surface, PDO is their anarchic opposite. PDO (which is not an abbreviation for anything) is what you get when you add a team's save percentage and it's shooting percentage. What you get is a number that's important precisely because it doesn't matter.

Teams actually tend to have relatively little control over their PDO; while individual players may be better finishers, or teams may suffer dips due to bad luck or injuries, over time the percentages will usually end up adding up to 100. So if a team has a PDO that's too much higher or lower than 100, say 102 or 98, then we'd generally expect that team to receive a visit from the regression dragons.

Take, for example, the Anaheim Ducks. Last season, after 28 games, the Ducks were in the bottom 10 in points, trailing teams like the Winnipeg Jets and leading the then-last place Columbus Blue Jackets by just four points. From that point, they were the second best team in hockey, accumulating 76 points and easily making the playoffs. And all it took to make the difference was a PDO swing. In the first 28 games, the Ducks had the fifth-worst PDO in the NHL (98.2); after that, they jumped to fourth (101.3). For the season, that evened out to a 100.3 PDO, right about what we'd expect.

So if your team gets off to a slow start, or if you want to throw cold water on an surprising playoff contender, PDO is the place to go. You can also get on-ice PDO for players, which will be useful the next time some fool on Twitter tries to throw plus/minus at you.

Zone Starts

The more research statisticians have done on Corsi and Fenwick, the more they've found ways to isolate other factors that may affect a player's possession numbers. Perhaps the most important of these is zone starts.

A Zone Start simply refers to the location where a faceoff takes place. If it's on the side of the ice with the opponent's goal, it's an Offensive Zone Start; if it's on your side, it's a Defensive Zone Start. On a team level, Offensive Zone Start Percentage will hew pretty closely to Corsi and Fenwick, to give you an idea of which teams are controlling the puck.

Query Results Table

			Corsi (EV)	Fenwick (EV)	Zone Starts (EV)
Rk	Tm	Season	CF%	FF%	oZS%
1	LAK	2015-16	56.1	55.9	55.7
2	DET	2015-16	51.8	50.4	53.6
3	TBL	2015-16	51.8	51.3	52.5
4	CHI	2015-16	50.9	50.3	52.0
5	WPG	2015-16	51.5	51.3	51.7
6	NSH	2015-16	52.3	53.5	51.4
7	NYI	2015-16	49.3	49.4	51.4
8	CAR	2015-16	51.4	50.7	51.4
9	PIT	2015-16	52.8	52.9	51.4
10	WSH	2015-16	50.6	51.1	51.3

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

However, oZS% (the shorthand way of writing Offensive Zone Start Percentage, if you want to further confuse non-stat folks) may be even more useful for players than it is for teams. For example, let's use Corsi For Percentage to compare two players:

Query Results Table

					Corsi (EV)	
Rk	Player	Pos	Tm	Season	CF%	CF% rel
2	Tyler Toffoli	C	LAK	2015-16	59.3	4.6
49	Sidney Crosby	C	PIT	2015-16	55.1	3.3

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

So Corsi is telling me Tyler Toffoli is better than Sidney Crosby? Get out of your mom's basement and actually watch the games, Corsi! Of course, it's not that simple, because I left off a key bit of context, their oZS%:

Query Results Table

					Corsi (EV)		Zone Starts (EV)	
Rk	**Player**	**Pos**	**Tm**	**Season**	**CF%**	**CF% rel**	**oZS%**	**dZS%**
2	**Tyler Toffoli**	C	LAK	2015-16	59.3	4.6	58.2	41.8
49	**Sidney Crosby**	C	PIT	2015-16	55.1	3.3	54.7	45.3

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

With this important bit of context added in, it's clearer that Toffoli had the advantage in Corsi because his team used him in a more Corsi-friendly role than Crosby. While both players lean offensive, the Penguins also used Crosby in a substantial number of Defensive Zone Starts, which makes sense since he's one of the best players in the world. The Kings, on the other hand, tended to shy away from exposing Toffoli to more defensive situations. His team asked him to defend less, and thus he was on the ice for fewer Corsi Against events, boosting his overall CF%.

Expected Plus/Minus

Corsi, Fenwick, PDO, and Zone Starts are the classics, but researchers and statheads are uncovering new ways of looking at the game all the time. Hockey-Reference added one such newer method to the site last season: Expected Plus/Minus.

Expected Plus/Minus is a stat for players that uses our shot location data to measure, not just raw shots, but the quality of those shots. For every player, this system looks at where on the ice a shot takes place and compares it against the league-wide shooting percentage *from that spot*, in order to determine the probability of that shot going in. The shot is then added to a player's Expected Plus/Minus.

So what's this measuring? Basically, what would we expect a player's plus/minus to be, based on the quality of shots his team took and allowed when he was on the ice and removed from noise (like the relative "hotness" of his goalie or the random luck of a particularly bad shot going in).

It doesn't tell you whether or not the puck is going in. What it's telling you is whether that player is consistently getting to areas, or not, where there is a good chance of the puck going in. This is useful because whether or not the puck goes in can often be subject to more random variation than whether you're driving play to particular areas.

Here's the top 10 in plus/minus:

Skater Statistics Table

			Scoring
Rk	**Player**	**Tm**	**+/- ?**
1	Tyler Toffoli	LAK	35
2	Anze Kopitar	LAK	34
3	Brian Campbell	FLA	31
4	Chris Kunitz	PIT	29
5	Colton Parayko	STL	28
6	Evgeny Kuznetsov	WSH	27
7	James Neal	NSH	27
8	Olli Maatta	PIT	27
9	Milan Lucic	LAK	26
10	Ryan McDonagh	NYR	26

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

Based on that, it might be hard to guess what two teams you'd expect to make the Stanley Cup. Now here's the Top 10 in expected plus/minus:

Skater Advanced Statistics Table

Rk	**Player**	**Tm**	**E+/- ?**
1	Joe Thornton	SJS	21.7
2	Patric Hornqvist	PIT	20.3
3	Tomas Hertl	SJS	19.2
4	Brian Dumoulin	PIT	18.9
5	Joe Pavelski	SJS	18.4
6	Justin Braun	SJS	17.1
7	Carl Hagelin	TOT	17.0
8	Ryan Getzlaf	ANA	16.9

Rk	Player	Tm	E+/- ?
9	Kris Letang	PIT	16.6
10	Marc-Edouard Vlasic	SJS	16.5

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

It probably won't always be quite as clear a correlation, but expected plus/minus is good for seeing what players might be benefiting from good luck or underperforming due to bad luck.

Adjusted Points

So far, the stats we've looked at are good for analyzing current performance. If you're trying to figure out how players or teams are doing now and why they're doing that way, these stats provide the context that's desperately missing from raw goal and point totals. But what if you want to compare historic performance across eras?

This used to be quite challenging because the game has changed so much over time. In the 1980s, for example, teams regularly scored five goals in a game. In the last 10 years, that happened much less frequently. With these changing goal environments, it means that raw goals scored numbers may not accurately reflect a player's "true" scoring prowess.

That's why, in addition to actual goals, assists, and points, Hockey-Reference has adjusted goals, assists, and points. These stats neutralize the effects that roster size, schedule length, and scoring environment had on a player's numbers.

For example, here are the NHL's best goal-scorers of all time, using regular old goals scored:

NHL Leaders Table

Rank	Player	Years	G
1.	Wayne Gretzky*	1979-99	894
2.	Gordie Howe*	1946-80	801
3.	**Jaromir Jagr**	1990-16	749
4.	Brett Hull*	1986-06	741
5.	Marcel Dionne*	1971-89	731
6.	Phil Esposito*	1963-81	717
7.	Mike Gartner*	1979-98	708
8.	Mark Messier*	1979-04	694
9.	Steve Yzerman*	1983-06	692
10.	Mario Lemieux*	1984-06	690

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

What a wild coincidence that, in a league that's nearly one hundred years old, six of the ten best goal-scorers of all-time started their careers from 1979-1986. Of course, now that you're a stathead, you know to instinctively doubt a coincidence like this. Clearly, context is bumping some of these players' numbers while depressing others who played in earlier or later eras.

So here's the same thing, but with adjusted goals instead:

NHL Leaders Table

Rank	Player	Years	G/A
1.	Gordie Howe*	1946-80	925
2.	**Jaromir Jagr**	1990-16	822
3.	Wayne Gretzky*	1979-99	758
4.	Teemu Selanne	1992-14	741
5.	Brett Hull*	1986-06	738
6.	**Jarome Iginla**	1996-16	690
7.	Brendan Shanahan*	1987-09	672
8.	Phil Esposito*	1963-81	671
9.	Maurice Richard*	1942-60	653
10.	Luc Robitaille*	1986-06	652

Provided by Hockey-Reference.com: View Original Table
Generated 10/13/2016.

Blasphemy! But it turns out, if you adjust for the eras they played, Gordie Howe and the still-active Jaromir Jagr were significantly more prolific goal-scorers than Wayne Gretzky. Don't worry though, if you look at adjusted points, Gretzky is still the GOAT.

Like any sport, knowledge of advanced statistics isn't a prerequisite to enjoying the game and they aren't meant to give you perfect, definitive answers to knotty questions like "Who will win the Stanley Cup?" But if you want to better understand the game, these numbers can help add shading and context to what you're seeing when you watch.