

Dick Cramer, co-founder of STATS, Inc.

Interviewer: Brian Hall

February 22<sup>nd</sup>, 2022

Brian Hall (BH): Thank you for joining me tonight, Richard. What are your earliest memories of baseball?

Dick Cramer (DC): They were actually rather slow to develop considering how intense they got. I was unaware, for example, of the 1950 Phillies winning the championship when I was eight years old. Most kids notice that sort of thing. I noticed that I've always had a sort of affiliation with lists of things. My brother who is actually younger, he brought home these baseball books and I said, "Oh these are cool." Went to a game, that was no big deal. Then I started following baseball more and more. I was probably about twelve, thirteen or fourteen years old when I started getting really into it. I spent some money that I'd been left behind by my parents to buy a baseball game and that was even better but by that time I was really my schedule was sort of centered around Phillies Games. The addiction had hit.

BH: Addiction, you say, because I feel this is a common thing with people who love statistics. It's this idea of lists or organization.

DC: Yes

BH: What are some of the other things you were thinking about with lists and statistics and organizing beyond baseball when you were getting more into these things?

DC: Cars, this is automobiles, of course classic cars for me are 1930s or 1920s cars, not 1950s or 60s, so that's one thing that I make lists of cars. The first thing was a Rand McNally Atlas and underlining the names of capitals of each country and just thinking a little more here and stamp collecting that would be another big thing. That's a hobby which all the boys seemed to have at one point that has pretty well disappeared by now. Television changes a lot.

BH: You also said, I think in your book you mention, that your sister also helped you.

DC: Okay, this is later on. Yes, it was in graduate school actually which by that time my fascination with baseball was the center point of my life. I was training to be a chemist, and this is going to present a problem because I'm really not very good in the lab but my first exposures to computers were with the punch cards and so forth. There wasn't very much charming about that. But, what my sister did, she was a summer high school student at the time at Penn for promising high school students to learn programming. She didn't care much for programming but a fellow classmate at her had a program that play baseball and she went to the trouble of getting those punch cards and sending them to me. I got a listing that's more complicated than you would think because computers were such expensive items in those days. My group had an account to use the computer, but everything was billed. I could hardly use the public resource for my private entries. Anyway, it turned out that I discovered then later that I couldn't do anything with program because I didn't have the right computer. Later when I finished graduate school and started my first job. It was another disappearing company, Polaroid, they had a computer they used for various computer defense functions, and I discovered it and I got my program working and started tinkering with it and discovered that my future was going to be a lot more comfortable for me working in computers and working with them rather than working in a laboratory.

BH: How did you end up feeling like “Wow, this programming thing is something that really resonates with me?”

DC: I just enjoyed tinkering with my program which is necessary because the code that I received was too big to run on this Polaroid computer. So, I chopped out the parts that weren't really very interesting and in the course of doing this, I discovered that this is something I really enjoy and then I had an opportunity, professional opportunity, to get back to school as a post-doc and I was there taking well above my grade level as far as computer skills but they really needed an organic chemist, a PhD organic chemist working this project and so I moved back to Harvard and actually that project became the basis for a Nobel Prize by Elias Corey because he was basically using a computer to illustrate a new way of thinking about how to synthesize organic molecules. The computer became a very graphic illustration. This computer had graphics which in those days was very unusual. It was a LightPen which had something like a mouse and it also played computer games, which I saw actually in Brooklyn in a exhibit of infotechnology was running this PDP-1, the first. This computer was as important to accessibility of programming as the PC was later. It became possible to sit down in front of the computer and tinker with it. You still had to be an expert and had to be very lucky to have the access, which I did, but it was a different time, we are talking about the mid-1960s now.

BH: So was the PDP, when you were doing your fellowship was that the computer you were using at Harvard?

DC: That's right.

BH: I think you mentioned in your book that I guess related to this work, was it Bill Gates or Paul Allen...

DC: ...Yes

BH: ...that ended up having exposure to the computer at Harvard when they were there.

DC: I didn't know Bill or Paul at that time. I did meet Paul Allen later, much later. It was interesting because the next step after the PDP-1 was the PDP-10. Corey's money got this computer and this became the computer that Paul and Bill most famously used to begin offline officially to write the program that became the basis of Microsoft. Bill was an undergraduate then and Paul and that's of course a well known story.

BH: Yes. It's just interesting how universities have certain computers. Different people are sharing time and computation expense and you can only go around so much. It's being shared by different departments and different students, right?

DC: That was true. Yes, this was *the* Harvard research computer. Each of these machines was, I guess. The PDP-1 I used and PDP-10 that Bill and Paul used later after I'd left

BH: So I mean it's seems that to me at least you happened to be at the right place at the right time...

DC: Absolutely. I consider myself...

BH: ...at many different points in your career, right? So you happened to be a PhD chemist that knows a little about programming, right, so you join an interesting project?

DC: Exactly. To say that I'm lucky. It's happened over and over again.

BH: And I mean that the same, I think, I mean we'll see later in the story that there's a lot of flights cancelled and coincidences and things happening where partnerships blossom and years of beautiful work that ends up coming out of those partnerships.

DC: In some ways, a bigger sense, I was in positions where I could do these unusual things. In this era, if I with the same sets of skills and interests it wouldn't resonate nearly as well as it had for me. And of course not before there were computers.

BH: What struck me with the title of your book *When Big Data Was Small*. What was it like back when you using these computers that you have access to to use baseball data. What did baseball data and simulations look like back when you were punch carding them into the computer?

DC: Well the basic logic is pretty much the same. You roll a random number, you have the computer generate a random number and you compare it to a scale. Let's say batting average. Your scale is a thousand. You can say if my random number is less than three hundred and that's the batting average of the player who is currently at bat, he'll get a hit or not if the value's higher. Of course there's a lot of sub probabilities of what kind of hit. You can look at frequencies of ground balls and we didn't have that sort of data at that time. What we had at that time were official baseball guides which wouldn't come out until spring of the following year. That was the only place for example to know how many walks a batter had obtained. You would see each Sunday a list of the top batters by batting average. Typically, the old triple crown stats. The homeruns and RBIs. The Win/Loss record for pitchers and ERA but that's about the only things you would know for a player and, of course, model.

BH: So when you were doing your simulations with the computers, right, you ended up stumbling upon some new baseball statistics?

DC: Yes, yes. My own version of the OPS. It turns out, I'll just mention it, it turns out if you multiply on base average times slugging percentage, instead of adding them as Pete Palmer has pioneered, you have some nice properties. For example, I guess the nicest is that product is exactly equal to or closely equal to the runs that player contributes to for plate appearance and there's more that follows from that. That's just sort of privy. They are equally accurate at representing the number of say...the thing you measure against is the only thing you have is the number of runs a team scores as of course some combination of how well the individual players do. So anyway, that's the general idea I think but back to the train of thought.

BH: A lot of this you were just doing individually, right?

DC: Oh absolutely.

BH: Then you did hear about an organization called SABR. How did you end up hearing about SABR and I guess later meeting Pete where you could bounce ideas off of and talk about analytics?

DC: Yeah. It was a little squib in an issue of Sporting News and all it said something about the Society of American Baseball Research and gave I guess the mailing address and I said I have no idea what these people do but I certainly feel like I should find out. So, I wrote the organizer from SABR, L. Robert Davids.

By day he was an imminent Washington bureaucrat and I think that describes his personal self pretty well but very capable. I told him what I was interested in, and he said, "Well, I don't know about that but there's this fella by the name of Pete Palmer who you might want to get in touch with." So, I did and Pete was, I would say, a little bit ahead of me in my thinking. He had come up with a, he had asked a question that hadn't occurred to me. What is the relationship between team's run scoring allowed and win/loss records? And Pete has an excellent the linear weights is what it's come to be called now, and I've become a convert ever since to using this as my way of thinking about that question. What we now called the Pythagorean theorem cause Bill [James] has a slightly different formula. I like Pete's.

BH: It was the beginning of you joining, what we would call now, the sabermetricians, a small group of people back in the day. It was a subset of SABR.

DC: Very subset. Bob did not think that sabermetricians should worry about these things. He was a little skeptical of them. He didn't think his readers liked it so actually I challenged him, and I did a user survey and Bob was entirely right. Very many of his readers were hostile toward the notion of analyzing baseball statistics. So Pete and I, all through the 70s, in the 80s, Pete and me writing letters back and forth. Pete reached out, more than I do...By nature I work by myself. But there were a couple of research ideas that I had which had been put in a book that Pete did with John Thorne, *the Hidden Game*. One of them was challenging the idea that clutch hitting is the skill of a player, I mean clutch hits happen, but is it the skill of a player that he can do this repeatedly? And the answer at that time there were only two seasons of play-by-play data in the world, something that existed but you could compare the records of the clutchness of a player in year one to year two. It was clear in just those two years that there was no correlation. Now that we have RecordSheet we can say the same thing at a much larger scale. Somebody like [David] Ortiz, our new Hall of Famer, he definitely hit some of the most clutch seasons of any player in history but if you average over his whole history there's nothing remarkable about it.

BH: I've read your work on this and I understand like random noise, right, it's like it doesn't hold up over time.

DC: Random fluctuations.

BH: Yeah.

DC: Yes. Exactly.

BH: But what do I tell somebody in New York City where I am now? Jeter is Mr. October. He is the most clutch guy ever in the playoffs. How do I reason with them?

DC: You can look at the record. Now there's lots of people these days, especially, who don't like to look at the record. I can see it myself, for example, let's say I'm doing, playing a simulation game myself. I feel my team is hot. I feel that the team is more likely to win the next game and even as a rational human being, having looked at all the data, I know that's not true. Human nature is what it is. So I wish you luck in your conversations but well, there are an awful lot of people trying to get people to listen to an awful lot of things these days. They don't want to hear.

BH: So going back to this, I guess. The time you and a slight group of people were very interested in looking at baseball data and maybe testing hypotheses with data and seeing maybe what the truth of the data would show. The fans back then did not view baseball statistics in this manner. What drove you

to keep doing the work you did, even knowing that there was resistance even from passionate lovers of the game, like SABR members, what kept you and your crew going?

DC: Well, I felt I was learning something. I think that's the characteristic of the scientist, I think, to be curious and feel like you're learning something. I guess the answer is it's good in itself, from my point of view. I don't have convince anybody else to take pleasure in what I'm doing for the most part. Yeah, praise is good but it's not what I'm thinking of when I'm doing the work, the possibility of praise.

BH: I mean you are primarily a scientist, right?

DC: Yes.

BH: And you did a lot of amazing things. I think lot of scientists are driven by the scientific method, advancing knowledge and research for humanity.

DC: No. I think that for the most part. Actually, I worked at a pharmaceutical company where you'd expect to hear that if people talked about it a lot. They really don't. They are paying their mortgage, they are doing activities. In my case, well my dad was a chemist and I was pretty good at it. I kind of enjoyed it but you usually don't discover...Most scientists are not fortunate enough to discover things. It's either not in them or they are not...I was in the right situation over and over again in my life and my best description of myself in one adjective is lucky.

BH: Lucky and you worked very hard, right? The amount of programming languages and hardware that you were working with over the decades, it's really phenomenal that you kept reinventing yourself and using the newest and latest.

DC: And that would be a lot of work. That's sort of because well this is stepping a little ahead in the story to the next chapter but when I started the communication from computers, these days is turbo, you and I are having no trouble, for example, if you want to write a program that works on the web there's a million ways and things you can just plug in. In those days, all you had was telephone lines and modem and even that was a bit of a luxury. It was slow and when ever, obviously we've moved up from the modem, but each time that happened it would be new technology and I would have to struggle with something I really didn't have the proper background for. It wouldn't be very well documented and the customers, in this case the baseball clubs, the Yankees, the White Sox, the A's, and the Astros were expecting of me to do. I would spend the baseball part, would be the fun part of the programming, and the communication would be the sort of masochism, I guess.

BH: Let's go back to how all this came about, right? You were having fun with doing your own thing, talking with others in SABR, occasionally sharing work, right? I think there was a quota with how many statistical articles [you could publish]...

DC: ...Yes. Yes...

BH: ...But then, I guess, things started to become more serious...

DC: ...Two factors I think. One is Bill James. Bill can write like nobody's business. I knew many of the things Bill wrote about but Bill wrote about them so powerfully and stubbornly and I think Bill James is one factor. The other was the personal computer. There was a lot of latent interest. There were a lot of people who wanted to do the kinds of things Pete and I had done if they had the kind of access that Pete

and I had to computers that could do that I kind of stuff. Once the personal computer became available, sort of alluded to when I talked about the PDP 1, this was a much broader scale anybody who cared at all about programming could give it a try. If they thought they might like it was easy to give it a try.

BH: Is this the time when the punch cards ended up getting discard and programming is more expressive?

DC: All of those things but it's obviously a lot easier to work with a keyboard with a program that's in front of you that you can change and see what happens than changing a punch card, giving it to an operator, and coming back in a few days to see if that worked. When you have an expressive media you can do things with it.

BH: A generation of people ended up getting access to personal computers and essentially experimenting and trying new things, right?

DC: That's right. That's what made the big change, that made analytics change, in not very much time from being a pariah to an emblem.

BH: But at this time, you know, data is still small. There's not play-by-play. It's hard to get the type of data we are used to today. Could you talk about the kind of next phase about how you personally got more serious, how you got connected to people, ending up turning it into a business that you did for fun on the side but it was a business?

DC: Yes. Pete was instrumental. As I say, Pete's much more good about staying in touch with other workers than I am. He knew of a guy named Steve Mann who also lived in Philadelphia as I did then. And Steve was working with the Houston Astros, he was sorta of a proto-sabermetrician. Craig Wright, both in Texas, both were Texans. There was a guy, trying to think of the exact sequence, it's exact in my book. In general, Steve...

BH: ...There was some sort of conference you were at and the flights ended up being cancelled...

DC: ...the flights, oh!

BH:...And you ended up collaborating more, right? That was part of the story.

DC: ...Well. Steve had tried to get a sports marketing firm called Specific Select in San Francisco interested in working with him and the Astros. They did not hit it off. I sort of said, I'll work with you. Steve was not a programmer. I said, "I'll see whether I can write a program that can record play-by-play in the progress of a baseball game." And I had a little basic interpreter on the computer I was using for my work. It was a basic program. Of course, you need some examples of play-by-play. The only thing I could find was the book that had play-by-play of the World Series. If you took a look at the first game of that World Series you will see that in the first inning...I thought that was a pretty good challenge for what I developed. Steve and the guy in San Francisco did not hit it off. I had a meeting near San Francisco, in the [inaudible], and I was about ready to head home and the flight got cancelled so I said, "I'll give this guy, Matt, a call because I got a day here to kill." And he said, "Can you program?" "Yeah, that's my role." He said, "Let's go to a computer store." Computer store? What the heck is that? That would be like a big box Best Buy as far as I understood so we went down to the computer store and they had...This was really the first time I'd seen a personal computer. They showed me the Apple II. How you could plug it in, sliding cards into slots by opening the case up, sounded pretty delicate to me. Either

way, bottom line was, I thought this could be sort of fun to do...Coming back to why this happened. The marketing firm had just persuaded the Haas family to buy the Oakland Athletics from Charles Finley. At that point, Charles Finley was, the club had slipped so badly that its' broadcasts were being carried by college shows and stations. So they really wanted to spice up the As or something new so that was the project that Steve...They thought something sexy had just happened down from the marketing firm premises and that was the Apple company. They thought 'Wouldn't it be cool to get Apple involved in broadcasting? Something like if Ricky Henderson came up, it would say something like, "His fifty-first passes the all-time As record" or whatever it might be. Or "Look Dwayne Murphy is not hitting lefthanders at all." There was no right versus left kinds of information, that sort of thing simply didn't exist. And could I then write a program that would do these sorts of things that was what I was doing when I was writing that little basic program. The basic answer was that I could do that but I needed an Apple II so to start I got in an airplane and came. I sort of outlined how I thought it should work. The big computer, I was using to do the main work, should do the heavy lifting and the Apple should be used to input the game and simultaneously display stuff that the broadcaster could use and that's what I wrote. We had some helping hands from Apple at that time in terms of some of the technical things. Again, it sounds real easy when I talk about it but given what was actually available it was a challenge, especially since I had only occasional telephone support to help. And there were times were I sort of tore my head. I'd go to bed and I'd think "I'm not going to solve this" but in the morning I'd get up and figure out what I could do. So anyway to make a long story short, by mid-Spring Training I had something that I could take down to the ballpark and the As and Cubs were playing, Hohokam Park. I took it out to the field and discovered certain things wouldn't work and I discovered certain things would work. It was ready for Opening Day. At least it seemed to be. I remember the As got off to a terrific start that year. They were like 19-3. But there were problems. For example, this program which we called Play Ball never had a control z, a backup, that's sort of problematic. In its entire history if you wanted to change a game, you'd have to...later on you could edit it but never during a game could you back up and say that play was wrong. You would have to edit after the game. There were bugs. It turned out to be possible to have runners disappear on the same base. And of course, I never did hear directly how the announcers felt about this display monitor sitting next to them. They didn't complain, I know that. Then it had to be taken on the road and somebody had to go with the computer to set it up and to enter the game and go through all that part of it. The guy that did it became a...with Colorado had a pretty extensive front office experience. Fortunately, before the walls came crashing down, we hit a strike during the middle of the 1981 season. By the time the strike was settled, I'd fixed the worst holes and things continued. That Fall...

BH: So I just want to understand how the play-by-play is going in. Is somebody watching the game and typing in each play?

DC: Of course. Let me say, there's a single person. The display screen is 80 characters by twenty-four line screen, eighty characters wide. I laid it out so that at the top of the screen there would be the notes. The stuff about how this guy is hitting against left-handers. At the bottom of the screen would be the place where the operator would be entering in the pitch; ball, strike. It's a ground ball to shortstop, what's the play? It's a 6-3. There's a runner on first, where did he go? There's a notation for doing all these things that you type in at the bottom. This is important too to make sure the notes at the top of the screen are referring to the player who is currently at the plate and not the guy who's going to come up next or the guy who just came up. With this design, those two things are synchronized on the screen at the same

time. So there would be this guy who is entering the game as it goes on. Of course, there's two displays on the same feed. One of those displays would sit in front of the broadcasters, I suppose they used it. The other would be the guy who is entering the game. Does that make sense?

BH: Yes. And once that information is in there, what were the A's doing with that information. What were the announcers supposed to be doing? What was the product?

DC: The product was only these notes. At the end of the game, there was a program that called up, it was the PVP-10 I was using, a program that Pete Palmer very kindly wrote for me for this project would generate a file of notes. Basically, the splits. Basically, it would be that kind of thing.

BH: So the splits wouldn't be calculated in the middle of the game. You'd wait until the end of the game to calculate the splits?

DC: Yes, exactly. The standard lines of this would be updated. But to do more than that, the program would need reference one...Because at the start of the game you would know the player's stats lines. Okay, he came up, bats increases by one, got a hit. He scored a run. Those things would show in his stat line when he came up next but the information about what he had done against lefties or righties was available only by going back to the big computer where it was being stored.

BH: So at the end of the game there's these notes that are generated. What is the team or what were the announcers supposed to do with those notes after the game?

DC: They'd use them during the game to be able to say how Armas is doing against right-handed pitching. It was not used for any other function at that point. Now come the winter, the sports marketing firm I was working with thought that other clubs might be interested and there was nothing in the system tied it to the As. So they talked first to the White Sox. I had a very nice association with the White Sox over the next few years. The White Sox wanted to use it, let's see, Tony LaRussa is there and particularly Dave Duncan thought he saw real promise if we would add pitch type and location to the information about the pitches. That was easy. But also I had to make the program communication run faster and that was hard.

BH: Were you aware that the A's were using any of these splits, let's say the lefty to righty split, to inform their managerial decisions with respect to their players or was this just an Apple announcer type of partnership?

DC: Billy Martin was their manager.

BH: When other teams potentially became interested, were they interested in using it for similar reasons or were they thinking about using it for their own managerial decisions.

DC: For the White Sox it was for their own managerial decisions and in particular, the champion, when you sell something to a company they are using an internal champion. Jack Gould, he was a close, longtime friend of Jerry Reinsdorf, still the White Sox owner. He was handling player salary negotiations for the White Sox and he felt that having this sort of information to take into an arbitration hearing would be a real asset to the club. He said there were several uses going forward. The use of the program to help on field decision making and inform the front office about player skills and weaknesses and especially in negotiating salaries. This is what it was for. And another thing, sort of interesting sideline thing, they also wanted to put all their scouting reports on a computer. I set that up for them as well.

That had its own challenges because the way you would typically do that in a database management system the data don't exist for...That's a sideline point of the league.

BH: So today are you more of an Apple person or PC person?

DC: Oh PCs. It was very clear...I'm using Mac as we talk but it's a matter of going where the capabilities are.

BH: The hardware.

DC: Exactly. In fact, there was a big transition began part of the overall history of this whole enterprise in going from Apple's to PCs and using our own computers, STATS own computers instead of the PVP-10 to do the heavy lifting.

BH: Now you have APIs so you don't have to worry about computers communicating?

DC: Exactly. That's about three or four generations ahead of where we were then.

BH: I know you were doing this programming alone but you also had several people involved with this business. Can you explain who they were and how they were involved?

DC: Yeah. Their core business was putting asses in seats. They would work with companies with sport franchises of all kinds to try to improve attendance. So, there was one of the programs should be bug free, that involves testing. That was something that I must say I'm not very...that's like lab work for me. Anyway, just creating the programs was since this...I should emphasize this was a sideline. All this was a sideline. I had a full-time career by this time, tying computers to drug discovery. If you asked a random sample of people around the world, they would know me far better for that work than for anything I've done for baseball. So obviously there's an upper limit to how much time I could spend. I didn't really get paid very much.

BH: Who were the other people you were working with? You had a career and you ended up doing this on the side for fun or for your own enjoyment, right?

DC: I got a slice of it. We set up this company called STATS Inc, which is still around today. As I said, my salaries, with my wife's, are more than enough for my needs. But the idea owning a bit of a company is something that sounded cool to me so that's how STATS, the company that's still obviously doing very well today got started. There were three names I can mention. Matt Levine who owned, he was an alumni of McKensey, the company in New York, I'm probably flubbing their name. Don Leopold, who was doing the testing. And the guy who worked for the clubs directly, his name will escape me, but Tom Black. He came up with the name. He was the person who got the cases for the Apple equipment, the equipment cases so it could be moved. He did training. He was on site in San Francisco. I saw almost nothing of the A's but I saw quite a bit of the White Sox, the Yankees and Astros later on.

BH: So, you had the initial business founders and partners at STATS Inc. and you were doing work with the A's and White Sox. What other type of work did you do with STATS Inc?

DC: Basically my role was that of programmer, the technical guy. And there were two other clubs that I worked with. My San Francisco colleagues, Tom Black and Matt Levine, did all the sales and pricing and so forth. And I gave them a good deal because I was enjoying myself. I owned part of Stats Inc. whatever that was at that time seemed that was perfectly satisfactory financially and a couple times, two other

clubs that I worked with. I had no contact with Oakland. I had quite a bit of contact with the White Sox. That was a very friendly and personal relationship I had with several people all the way up to Jerry Reinsdorf, the President, where I embarrassed myself by shorting out his Apple II when I tried to stick in the card. The other two teams were the Yankees, where I worked with Doug Melvin, who later became a pretty well-known general manager with the Braves and Texas. And in Houston I worked with a guy, well I think Nolan Ryan was the instigator of Houston getting the system. But the fellow I worked with was a guy who worked for NASA, a doctor who also consulted for the Astros. He did a pretty good job with it. There were funny stories that I won't take the time to go into about my interactions with the Houston Astros. So those were the four clubs that I dealt with. Also the Columbus Clippers, who were a Yankees farm system at the time. I went to the Yankees pretty regularly. I was living in Philadelphia for a lot of that time. I went to Houston a couple of times.

BH: So how did the Nolan Ryan end up getting behind the system? Why do you think he found it interesting?

DC: I don't know. Nolan Ryan was a sort of semi-god in Houston at that time. Whatever Nolan Ryan wanted, he got. Also, it was extremely strange when I visited Houston for the first time. I'd set up the computer the first time I went and this was their general manager Alan Rosen. He had no idea the team had a computer. So clearly somebody else was behind it and Nolan Ryan's name was mentioned. I suppose that Nolan just thought that...they were very interested in pitch type and location information which by that time had started to be recorded. The White Sox had wanted done so I'd added that to the system and so I suppose that's what he looked at but again I didn't have close interaction with Nolan. I'm just inferring that clearly the front office wasn't directly involved because they didn't even know it existed.

BH: So, what do you think they were doing with the pitch type and location information, like the teams?

DC: There was, I can't resist this story. When they stopped and said, "The only thing they did convince Mike Scott that his curveball was no good." If you know Mike Scott's subsequent career. He started throwing a palm ball, I think that's what he called it. He had a couple Cy Young level performances after that. But at that point he was not an asset pitcher and suddenly became an ace so in my view the Astros, if that's all they got out of it, they got a pretty good deal.

BH: But you ended up later expanding to different clients who were primarily not teams. It was for media clients?

DC: Yeah. It was this point there's a break. The people in San Francisco, the parent company, went broke. Matt Levine did not want to lose confidence in the sports world in general, that was where he was based. So after some negotiation I agreed to simply take on the business so I convinced (*ineligible*) to sell it to me but I was willing to continue the arrangements. Which continued with the White Sox and Yankees for several years. So now I'm the sole proprietor of this thing and I'm looking around for partners. At that point, I was still club supported. Those were the only customers. During that time, I cast around for people to work with. I called Bill James and he told me about these fellas who were doing Project Scoresheet in Chicago. I drove up there and John Dewan and I meet, and our wives meet, and we immediately hit it off. It was apparent that the programs I had put in a large enough computer could really remove a lot of the steps that they have in terms of turning scorecards into computer readable stuff. So, there was a year of sort of semi collaboration. And then John had a huge falling out

with Project Scoresheet. As a result, here's when the company got reorganized with me putting in some money in the existing business and John putting in something comparable. There were a couple others. Bill James was a very minor investor. And with that money, the reason we incorporated was so we could buy our own computer. There was a new computer, a small computer, an affordable computer that didn't need an air-conditioned room to run but could still do run up type programs if you'd like, it was called a MicroVAX Digital. That money was put together, so we used that to buy a MicroVAX. At that time we still needed to run the business so I agreed with the company I was then working for for software, the one I was drawing my paycheck, I agree with them if they would answer the phones on behalf of Stats I would let them use my new computer which would be an asset to them. I didn't need all its capacity all the time. So, STATS didn't need it and customers didn't need it. There was a massive rewriting program at that time with the Apple II. I actually brought an Apple 2 with me, and I can pick it up as a souvenir so we moved on to the PC. John and Sue then set about building the business. The business was clearly going to be to collect all the scoresheets and the customers now. Now we have the business and what are customers are going to be. The first customers, we knew what we could do because John had done already done it with Project Scoresheet write a book of stats. Because of at that time, you could get the season ending stat book completely only next Spring when the Sporting News would come out with it's annual guide. There were also issues of a self-publishing program which John Dewan mastered. We had official stats coming out just a few weeks after the season ended. Just a week or so. That was clearly something we knew that we could sell. Another thing that was getting underway was fantasy baseball. John and I were slightly a little bit contemptuous of it, but our first employee Carmen Corica was a real fan and Bill James wanted to design a game, so I programmed it. You could play Bill James' Fantasy Baseball online. John went on from there. As we gained credibility there was a start up in the newspaper business that thought they could use our stats. They winded up going with the established player but USA Today was impressed enough with the potential competition that they thought they should greatly enrich their baseball product. So, they could do something for us that we could not do; get into Major League press boxes. John and Sue did an extraordinary organizational job. They got programmers, they got the scorers and everybody out there in time. It was a monumental feat to start. Once we were in USA Today, we became credible. We then got a contract with ESPN. Another year later Associated Press said, "it's silly for us to spend time doing it if these guys can do it." I'd written all the programs for a while I had the sort of ego trip of thinking "Boy, my code has generated about as many lines of press as anywhere in the world." That was fun. I'm summarizing several years obviously of development and John found other customers like game companies and helping other people with their fantasy games. Originally, of course, just getting daily splits, lefty vs. righty, we were the only game in town for that if you wanted that for whatever reason. That's a lot of history there to answer your question. That's how things went.

BH: So, what appealed to these traditional media outlets about computers, digital data, the internet? How did your program and data kind of update their products?

DC: It was a sort of arms race. It was clear that people wanted more sports information. That fantasy baseball was not an oddball. It was something people loved. For a while it felt like being the only arms dealer in a war or well the only supplier, rather, in a war. In the case of the Associated Press, we just made easy what they were already doing. They didn't have to pay stringers at the ballpark to gather the information we could gather, that our scorers were gathering. In the case of ESPN, actually we had an employee who worked on site at ESPN who could use the database and could answer questions. The

first time I went up and demo mode, well they had a game of the week. They had a game every night, if I recall correctly. The first time I went up as a demo project, Dickie Thon of the Phillies hit a home run and I could quickly tell them the last time Dickie Thon hit a home run in the 9<sup>th</sup> to win a game and all sorts of peripheral information. Again, this is run of the mill stuff now, but it simply didn't exist at that time. People liked it. They may go overboard now but I guess that's a matter of taste.

BH: What year was it that you ended up getting the USA Today contract? Was it before or after the strike?

DC: Oh, it was before the strike. This is all before the strike. The big strike of 94. The big strike was a very difficult time for STATS Inc. My guess by that time several more things had happened. John had several initiatives that I happily supported. STATS started gathering information on other sports. Football. And again, I did most of the programming. And then John could see how fast the company was growing. John has the normal ambitions of any businessman to grow his company as fast as he can. At that time the only added capital was the bit of money I had added in at some point to buy another piece of hardware but this self-financing wasn't going to build the business, so we had gotten an outside investor who was famous in his own right. That's Paul Allen of Bill Gates and Paul Allen. And then for a lot of reasons, none of which in retrospect mean very much, John and I started getting quite mistrustful of one another and this situation where we did not agree with how STATS as a company should respond to the challenge of the strike. There's no sense going into in detail, but the bottom line was that I left the company, was ejected from the company, if you like. At this point John and I are good friends and I participated in a Zoom with him only a couple weeks ago so that's just a sort of passing interest at this point.

BH: The strike was a difficult time for STATS. I think Paul Allen ended up investing before the strike, right?

DC: Yes. From his point of view...this is a little complicated. At that time, the World Wide Web was just becoming a phenomenon. And as I understand it, he had organized a company that was going to develop systems that other companies could use to bring World Wide Web in-house, and he had developed a contract with ESPN to do that. He invested in STATS as his supplier to that division. So, it was basically he was investing to help support this other business which had much broader ambitions.

BH: But then the strike happened and how did it effect STATS? How close were you guys to...

DC: ...To bankruptcy?

BH: Yeah.

DC: Eventually John realized that this could go on longer than...I mean we did have the outside investments so we could run head on. It was painful but the growth was being invested in survival. Eventually John agreed that the company should be sold. We put out to bidders. And here things get complicated and ugly. The bottom line though was at the end of some negotiations that I didn't actually participate very much in because I'd been ejected, STATS became an independent entity. Or it continued as an independent entity which is today it's no longer the incorporation that I originally organized. That's a few years ago now. That was more than twenty-five years ago.

BH: So, you're saying that STATS ended up being purchased...

DC: No, it became independent. ESPN put in a bid for it, and we rejected the bid. I agreed to reject the bid. I was the dominant shareholder still. STATS had its own business trajectory. At that point it was still independent entity for four or five years under John's leadership again continuing.

BH: So, I guess eventually John did sell it?

DC: Yes. That was quite a few years later. That was in the peak of the dot-com bubble. He got an excellent price for it. I remember thinking that's more than the last time the Phillies changed hands which was twenty years before that. But still. At that time, for various reasons, I didn't hold anymore stock or not much. One thing I have not been in the course of my career is a good investor.

BH: I guess I want to trace the path of the company because your legacy code, these ideas that you had...

DC: Yes.

BH: It's not like these just go away when. Lot's of people have starts ups and the company does not survive, the ideas don't survive, but this company that you did found survived and the ideas have continued to this day, right, and now are almost common place, the kind of information you guys pioneered.

DC: I would say I've been a very lucky man to be in the right place at the right time and to go in the right direction and that's true of my other work, as well. I mean my paycheck job. It's certainly a great joy to think about it at this point in my life.

BH: Well, It's about being in the right place at the right time and also being able to do very complex problems and a lot of work related to...

DC: And that's good luck too, isn't it?

BH: That's very true. There's a very good feeling when you finish a program and you think it's...

DC: ...Absolutely

BH:...long way off with the program and it works and you're very happy. You get a high.

DC: There's some bumps. By the time you get to find all your mistakes, or most of them, there's some kind of, as we all know, mistakes show up in code long after it's put in use and my code is no exception.

BH: So, after your involvement with STATS Inc. what was your relationship with baseball?

DC: I didn't have a relationship with baseball for about ten years. I got back but there was a meeting at SABR. I attended that meeting. Actually, I guess my only contact was that I enjoyed the trivia contest. I think I mentioned the last time we talked about I introduced trivia in the early meetings and it was not met with approval. In fact I'm tempted to go over to show you the bat I won, it's sitting in the corner just out of your line of sight for winning the National Trivia Contest one year, might have been that year.

BH: You were able to beat Pete Palmer in trivia?

DC: Yeah. The meeting was in Boston, and I looked out at the Fenway Park's scoreboard and I couldn't believe it. They were showing OPS, apparently this had been happening for sometime, and I was like

“Wow!” I had dreamed of that happening at some point there it is at Fenway and it’s on the scoreboard, OPS. And Pete also had been interviewing for a book. He’d introduced me to Michael Lewis, he just said he was a New York Times reporter. We talked a little bit and then I remembered I’d read some of his earlier books. And we had some actual discussions. They were more discussions than interviews after a while. Of course, I became an important person in Moneyball, in the book, not the movie. So, at that point, I noticed Retrosheet. Now it took a while for me to get as engaged with Retrosheet as much as I wanted to be but what Retrosheet was doing, under Dave Smith’s leadership, I started working there and that’s what I’ve been doing for the last ten years or so. I have a project of my own at this point. Dave didn’t think the Nineteenth Century was sufficiently well-organized baseball enough to be in Retrosheet’s scope so I started entering these seasons. These days getting a hold of newspapers is not a problem. I can work on the same computer that we are talking now and look at all the stories of all the games of each game simultaneously. So, on this project I started by completing recording Retrosheet style play-by-play for the 1899 season with the famous Cleveland Indians or Cleveland Spiders, rather, the greatest losing team of all time moving on to the 1900s. I asked Dave whether to move forward and he said no so I moved backward. Did 1898. Someone else had put their finger on 1896 and 1897. So I’ve completed the all-time hitting year of 1894 and I’m about a third of the way through 1895. I hope that these games will go into where all such stuff belongs Sean’s Baseball Reference online. Basically, it’s a primary source of play-by-play information. Again, I’ve been lucky. Again and again, I’ve been able to take advantage of new technology that seemed to me enjoyable and worth doing.

BH: What do you think appealed to you about the Retrosheet project?

DC: Well, I guess what’s always...For some reason baseball stats, they do interest a lot of people too, but it’s a collaborative project and people took it seriously and originally it was much harder because you had to get the newspaper stories and get people to clip stuff and send you stuff. Occasionally there was games in those days would be complete play-by-play, I guess more than half, but for other such games, I call it sort of baseball sudoku, where you have the runs scored, typically play-by-play from the game stories from those and you have a box score and you fill in the things you can fill in and you basically build a box score in a spreadsheet and then you start your scoresheet off to the right of that and you fill in what you can from the stories and from there you take your best guesses. You still have parts left over from the stories and the box score and you try to piece those in the right place. You have to make the line scores match. You have to make the right scores in the right innings, usually that’s part of the game stories. Sometimes there’s nothing. One good example is a game in which the Phillies got 36 hits which is the all time record in 1894. The game story said it’s just too complicated, so we aren’t going to tell you anything. You have to come up with a reasonable play-by-play to match the line score and the box score and it’s an interesting exercise. I’ve developed some techniques for doing that too but that’s a lot of work to come up with assembling 36 hits and 29 runs in 8 innings, I think it was, and getting hits and runs to come out to the right places according to the box score.

BH: So, I know Retrosheet has been around for a while. What do you think made it kind of hit a wall with certain dates? Where it couldn’t go back further in time and there’s only a few of you now working in that era?

DC: Well of course when Retrosheet started, when they have all play-by-play data, when I was with STATS, that’s the point at which complete play-by-play data exists in 1987, when STATS with John’s leadership and my programs collecting all that had been made available of that play-by-play information

in one form or another. Going back a few years, for example, Dave has got complete play-by-play got the newspaper reporters, I believe they maybe score sheets for every game he played in. I recall, for example, Cincinnati has a fairly completely play-by-play record, pieces of play-by-play from the late afternoon edition or something like that but otherwise it's the process I described. Where all you have is a box score and sometimes they disagree. The box score from a couple of papers and stories from a couple papers and it's a sudoku like process to fill in all the blanks in a sort of phantom score card and it has to be consistent with all that information you have. Obviously, it's not absolute truth but it's definitely the best you can be done today.

BH: Do you remember your reaction when you first saw all freely available baseball information that was very detailed on the internet?

DC: Well, actually I forgot to mention one of the projects that John wanted to do when we got the initial investment which was actually my problem. That was to make for the first time to have to have a play-by-play dynamic. At that point you could have play-by-play after the game was over, but we wanted to have that information accessible. John wanted to make that accessible live. Which for me that was a tremendous challenge making that. Again, at that time we didn't have the internet tools that everybody has today. I had to patch together some assembly code and some stuff. Well, I had to manage the process of keeping telephone line online and the ability to notice when it was offline, and the ability to retranslate the parts where it was missed. The internet was just coming...There were no browsers at this point. Browsers were another year or two into the future. I'm surprised I forgot that.

BH: Sounds like STATS was also involved with having the first live update database that was available on the World Wide Web.

DC: That's exactly true. Now that I think about it, it's not a pleasant memory because that code was almost impossible to debug. I'm sure it failed a lot of times for people, but I did the best I could.

BH: Well that explains why it didn't come to mind immediately, right?

DC: Yes

BH: A little more painful than other programs.

DC: Not altogether pleasant. As I say, I was a lucky man, and it has been that way all through my life. It's not like that now. It often feels there's a chance when there are breakthroughs that do things nobody could imagine could be done so I say I was lucky that these things did happen to me. But advice, being perfectly honest, try to make a hobby out of it. Don't try to make it your livelihood. That's what I did.