



زمانی که برنامه‌نویسان کنترل داده‌ها را به مفهومی به نام بانک اطلاعاتی منتقل می‌کنند، کار برنامه‌نویسان در آن مقطع متوقف می‌شود و کاربران بانک اطلاعاتی کار خود را آغاز می‌کنند که این موضوع بدی نیست. بعد از گذشت سال‌های متمادی بانک‌های اطلاعاتی به گونه‌ای مهندسی شده‌اند تا مؤثرترین عملیات مدیریتی را روی داده‌ها به‌ویژه در ارتباط با اصول زیرساختی فراهم کنند. ساخت بانک‌های اطلاعاتی که در گذشته توسط یک برنامه‌نویس انجام می‌شد، نمونه‌ای از این موارد است. بانک‌های اطلاعاتی معمولاً به شیوه مؤثری از حافظه اصلی استفاده می‌کنند و ترفندهای خاصی برای تقسیم‌بندی حافظه دارند. اما پرسش این‌جا است که آینده بانک‌های اطلاعاتی به کجا خواهد رفت و از همه مهم‌تر برنامه‌نویسان را به کدامین وادی هدایت خواهد کرد؟

این مطلب یکی از مجموعه مقالات پرونده ویژه «آینده برنامه‌نویسی» است. برای دانلود کل پرونده ویژه [اینجا](#) کلیک کنید.

بانک‌های اطلاعاتی در مسیر هوشمندی

مبادله داده‌ها با اعداد غیر هوشمندی که به داخل بانک‌های اطلاعاتی وارد می‌شوند و اجازه می‌دهند فعالیت‌هایی را روی آن‌ها انجام دهند، اغلب راه‌حل هوشمندانه‌ای به شمار نمی‌روند. آن‌چنان که بانک‌های اطلاعاتی به سوی هوشمندی گام برمی‌دارند، داده‌های آن‌ها نیز به همان میزان رشد خواهند کرد. دلیلش این است که تمرکز بانک‌های اطلاعاتی آینده روی قابلیت‌هایی بیش از ذخیره‌سازی اعداد قرار خواهد داشت. بیش‌تر سیستم‌های بانک‌های اطلاعاتی از مدت‌ها قبل شروع به استفاده از موتورهای گزارشی پیچیده‌ای کرده‌اند که در اصطلاح رایج هوش تجاری نامیده می‌شوند.

Web Workers: [معماری و کاربرد Web Workers](#)

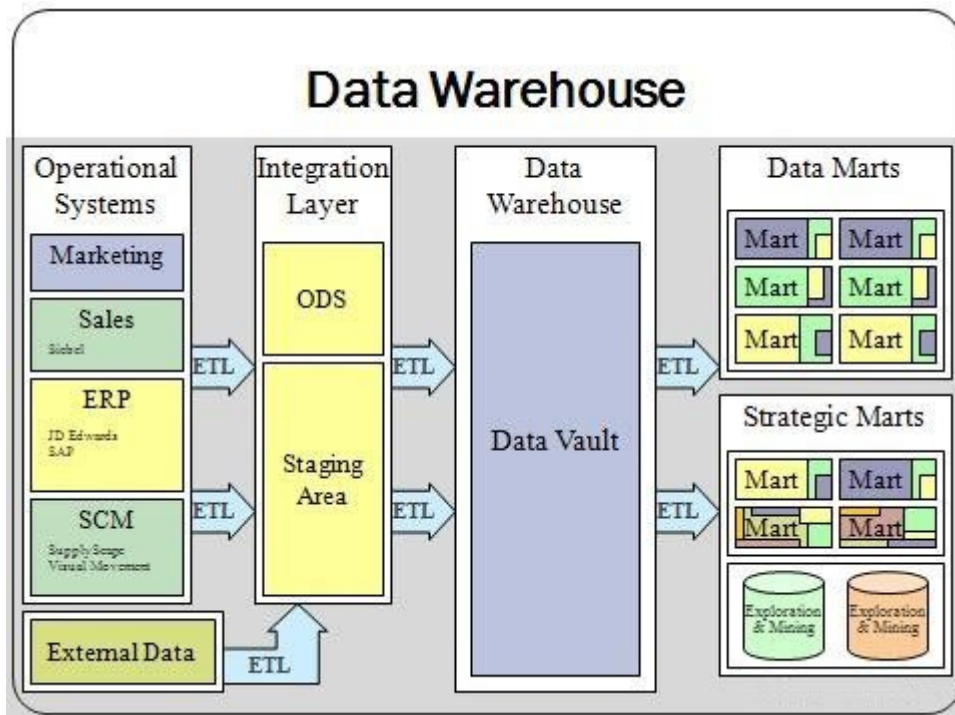
این ویژگی‌ها توانایی بیشتری در اختیار آن‌ها قرار خواهد داد، به طوری که بانک‌های اطلاعاتی را قادر می‌سازد تا الگوریتم‌های پیچیده‌ای را روی جداول اجرا و از الگوهای کارآمدتری برای جست‌وجوی داده‌ها استفاده کنند و خیلی از کارهایی را که امروزه «بزرگ داده» نامیده می‌شوند به راحتی انجام خواهند داد. استخراج اطلاعات از یک بانک اطلاعاتی و توزیع آن‌ها در قالب بزرگ داده‌ها به‌طور فزاینده‌ای وقت‌گیر و نیازمند برنامه‌نویسی فراوانی است. در آینده استخراج داده‌ها از بانک اطلاعاتی و اجازه دادن به موتورهای آن‌ها برای تجزیه و تحلیل به‌سرعت انجام خواهد شد. همین موضوع باعث کاهش برنامه‌نویسی به‌منظور استخراج مقادیری از داده‌های ذخیره شده خواهد شد. اما این مفاهیم نوین چه هستند و چه مزایایی را برای طراحان همراه می‌آورند؟

هوش تجاری (Business Intelligence) چیست؟

قبل از آنکه به ابزارهای مرتبط با بانک‌های اطلاعاتی نگاهی داشته باشیم، ابتدا بهتر است با مفاهیم هوش تجاری و بانک‌های اطلاعاتی تحلیلی که سبک زندگی بانک‌های اطلاعاتی را تغییر داده و بسیاری از وظایف برنامه‌نویسان را بر عهده گرفته‌اند، آشنا شویم. هوش تجاری مجموعه‌ای از تکنیک‌ها و ابزارها برای انتقال داده‌های خام به اطلاعات قابل فهم و مفید برای اهدافی در زمینه تجزیه و تحلیل‌های تجاری است. فناوری‌های BI این توانایی را دارند تا به مدیریت حجم زیادی از داده‌های غیرساخت یافته برای کمک به شناسایی، توسعه و ساخت فرصت‌های جدید راهبردی در کسب‌وکار پردازند. هدف از BI تفسیر حجم زیادی از داده‌ها است. شناسایی فرصت‌های جدید و پیاده‌سازی استراتژی‌های مؤثر بر مبنای بینش به دست آمده، به کسب‌وکارها اجازه می‌دهد از مزیت یک بازار رقابتی و بلندمدت پایدار استفاده کنند. هوش تجاری از تعداد زیادی مؤلفه ساخته می‌شود که از آن جمله می‌توان به ارائه گزارش‌های بی‌درنگ همراه با هشدارهای تحلیلی، بهینه‌سازی برچسب‌ها و استانداردها، روش ارتباط با منابع داده‌ای بدون ساختار و غیره اشاره کرد.

انبار داده (Data Warehouse) چیست؟

انبار داده سازمانی (Enterprise data Warehouse) سیستمی است که برای گزارش‌گیری و تحلیل داده‌ها به کار می‌رود. انبارهای داده‌ها مخازن مرکزی از داده‌های یک‌پارچه‌ای از یک یا چند منبع متفاوت هستند. آن‌ها داده‌های جاری و قبلی را که برای ایجاد گزارش‌های تحلیلی برای کارمندان سراسر سازمان ارسال می‌شود، ایجاد می‌کنند. از جمله این گزارش‌ها می‌توان به مقایسه سالانه و سه ماهه و روند تحلیلی مربوط به جزئیات فروش روزانه اشاره کرد. داده‌های ذخیره شده در انبار داده از طریق سیستم‌های عملیاتی (از قبیل بازارهای فروش به دست می‌آید) (شکل 1). انبار داده از سه لایه Staging، Integration و Presentation ساخته شده است.

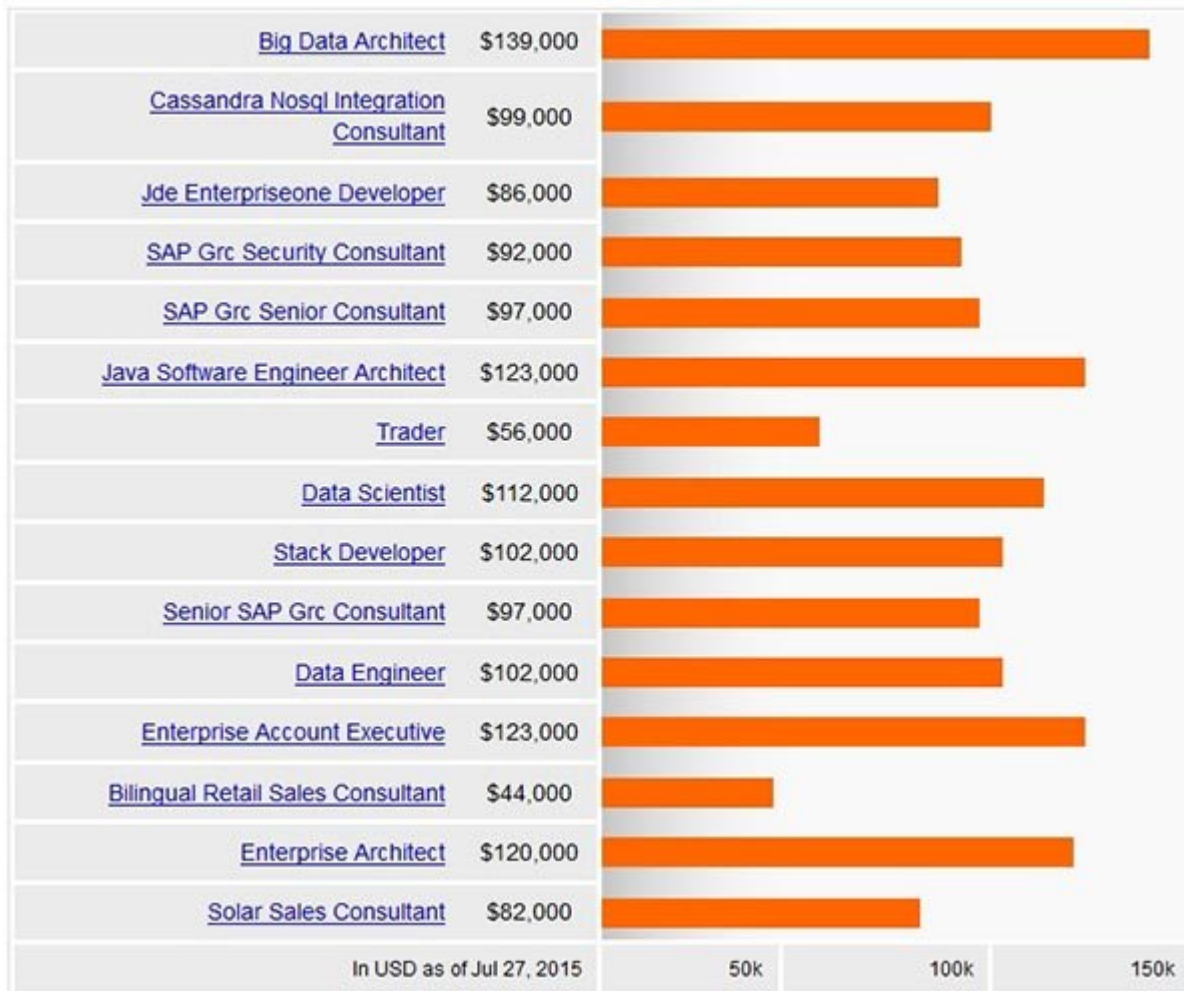


شکل 1: معماری انبار داده (Data Warehouse) شامل سیستم‌های عملیاتی، لایه یکپارچه‌سازی، انبار داده، و انبارهای داده.

اهمیت بزرگ‌داده

دنیای بانک‌های اطلاعاتی به اندازه‌ای گسترده و پهن‌تر است که با چند جلد کتاب هم نمی‌توان آن را توصیف کرد، اما یک حقیقت وجود دارد. در آینده حجم داده‌هایی که توسط سازمان‌ها تولید می‌شود، بسیار زیاد خواهد بود. به طوری که حجم داده‌های سازمانی هر پنج سال دو برابر می‌شود. حتی دستگاه‌های مختلفی که امروزه مورد استفاده قرار می‌دهید نیز داده‌های مختلفی را تولید می‌کنند. کافی است نگاهی به اطراف خود داشته باشید، گوشی‌های هوشمند، تلویزیون‌های هوشمند و جالب آن‌که هنوز دستگاه‌های اینترنت اشیا به طور جدی به میدان وارد نشده‌اند. این حجم از داده‌ها نه تنها به ابزارهای قدرتمندی نیاز دارند، بلکه به کارشناسانی نیاز خواهند داشت تا بتوانند این داده‌ها را مدیریت کنند و درست در این‌جا است که باید دید طراحان با چه ابزارهایی بهتر است حداقل آشنایی اولیه داشته

باشند. در آینده کمتر شرکتی نیازمند یک برنامه‌نویس خواهد بود تا بانک اطلاعاتی یا جداول مربوط به آن را برای آن‌ها ایجاد کند. روزگاری که عملیات ساده بانک اطلاعاتی در محیط بدون روح انجام می‌شد و نتایج محاوره‌ها در قالب یک نمایش ساده به اپراتور نشان داده می‌شدند، به سر رسیده است. به‌طور مثال، اگر یک معمار بزرگ داده باشید، می‌توانید به درآمد 139 هزار دلاری دست پیدا کنید. معمار بزرگ داده‌ها فردی است که مسئولیت معماری، طراحی و رهبری گروه برای پیاده‌سازی این موارد را در یک مجموعه از پلتفرم‌های تجزیه و تحلیل بی‌درنگ بر عهده داشته باشد (شکل 2).



شکل 2: میانگین حقوق سالانه برای 52 نقش مختلف در صنعت فناوری اطلاعات در سال 2015.

در این سیستم هوشمند، برنامه‌نویسی در کجای کار قرار دارد؟

اگر تصور می‌کنید با داشتن اطلاعاتی درباره با تی‌اس‌کی‌وال (TSQL) توانایی پیدا کردن شغلی خوب با دستمزد بالا را خواهید داشت، بهتر است تجدید نظر کنید. همان گونه که به اختصار گفتیم، تکنیک‌ها و دنیای بانک‌های اطلاعاتی به سمت و سوی هوشمندی گام برداشته‌اند، طبیعی است هرچه سیستم‌ها از فرآیندهای هوشمندانه‌ای استفاده کنند، کارهای سنتی و دستی گذشته مکانیزه شده که ابتدا سرعت و دقت انجام کارها را افزایش می‌دهند و دوم نیاز به تکنیک‌های برنامه‌نویسی که در گذشته پرطرفدار بودند را کاهش می‌دهد. طراحان مسلط به انبار داده و هوش تجاری به نسبت از درآمد بیشتری برخوردار هستند (شکل 3).

Average Salary of Jobs with Related Titles



شکل 3: میانگین دستمزد در ارتباط با عنوان‌های شغلی مرتبط با طراحی بانک اطلاعاتی

هرچند بیشترین دستمزد در ارتباط با طراح ارشد بانک اطلاعاتی قرار دارد، اما طراح ارشد بودن نیازمند مهارت‌های بسیار زیادی است، به طوری که برای هر یک از مواردی که یک شرکت به آن نیاز دارد، باید یک سابقه پنج تا دوازده ساله داشته باشد. اما طراح انبار داده لازم است تا از بینش مناسبی برخوردار و تجربه مرتبط با شغل انتخابی خود را داشته باشد. یک نمونه آگهی استخدام در ارتباط با طراح انبار داده را در شکل 4 مشاهده می‌کنید. در مجموع، میانگین دستمزدی که یک طراح بانک اطلاعاتی دریافت می‌کند، 92 هزار دلار در سال است (شکل 5).

Number of Openings:	2
Requisition Number:	15-10473
Job Title:	Data Warehouse Developer, Sr
Department:	17162-Knowledge Management
City:	Seattle
State:	Washington
Location:	70th & SPW Admin Offices
Secondary Location:	
Union:	None
Work Status:	Regular
Shift:	Day Shift
FTE/Hours per pay period:	1.0 FTE (80 hours per bi-weekly pay periods)
Hours Per Shift:	
Call Hours:	
Work Days:	Monday-Friday
Salary:	Depends on Experience
Job Description:	The Enterprise Data Warehouse (EDW) and Business Intelligence (BI) Developer, Senior supports the Hospital and Knowledge Management's vision, mission, and strategic initiatives by using technical expertise to design, create and develop automated EDW and BI solutions and by providing internal consultation and project management. The incumbent provides knowledge in business/clinical processes, analytical expertise, technical programming skills,

تعداد 4: «تعداد 2» در مورد این شغل. برای اطلاعات بیشتر، لطفاً به لینک زیر مراجعه کنید.



تعداد 5: «تعداد 60» در مورد این شغل. برای اطلاعات بیشتر، لطفاً به لینک زیر مراجعه کنید.

NoSQL قیام خود را از مدت‌ها قبل آغاز کرده

از آن‌جا که بانک‌های اطلاعاتی همچون اس‌کی‌و‌ال، اوراکل و MySQL سال‌ها است توسط جامعه حرفه‌ای مورد استفاده قرار می‌گیرد و برای همه خوانندگان شناخته شده است، ما از آن‌ها صرف نظر می‌کنیم و به‌طور مختصر به بررسی NoSQL خواهیم پرداخت. NoSQL را می‌توان راه‌کار مناسبی برای مدیریت داده‌ها دانست. مایکل استون بریگر، دانشمند علوم کامپیوتری، نامی آشنا در دنیای بانک‌های اطلاعاتی به شمار می‌رود. او کسی است که در طراحی محبوب‌ترین سیستم‌های بانک اطلاعاتی همچون PostgreSQL، Ingres و Vertica مشارکت داشته است. بریگر بر این باور است که در حال حاضر راه‌های به‌مراتب بهتری برای مدیریت داده‌ها نسبت به بانک‌های اطلاعاتی رابطه‌ای وجود دارد. او می‌گوید: «در میان شرکت‌های مختلف در جهان، فیس‌بوک با مشکل شدید مدیریت داده‌ها دست به گریبان است. آن‌ها از سال‌ها پیش در فکر جایگزینی برای MySQL بودند، اما هنوز نتوانسته‌اند ابزاری که بتواند این مقیاس از داده‌ها را سازمان‌دهی کند، پیدا کنند.» اما یک پرسش وجود دارد. اگر من به‌عنوان خواننده این مقاله بخواهم تمرکز خود را روی یادگیری برنامه‌نویسی بانک‌های اطلاعاتی NoSQL قرار دهم، نمی‌دانم کار خود را از کجا آغاز کنم. چه بانک‌های اطلاعاتی امروزی بر پایه NoSQL کار می‌کنند؟ برای آن‌که یک دانش اولیه در این زمینه داشته باشید، در این بخش شرح مختصری از بانک‌های اطلاعاتی مبتنی بر NoSQL را ارائه می‌دهیم.

رده‌بندی بانک‌های اطلاعاتی مبتنی بر NoSQL

پنج طبقه‌بندی مختلف در ارتباط با بانک‌های اطلاعاتی NoSQL وجود دارد. هرکدام از این دسته‌بندی‌ها با زیرگروه‌های مختلفی همراه می‌شوند. به دلیل ماهیت هم‌پوشان این گروه‌ها، ارائه نمونه‌های مختلفی از آن‌ها شبیه به بانک‌های اطلاعاتی رابطه‌ای به‌سختی امکان‌پذیر است، با وجود این، یک طبقه‌بندی عمومی بر مبنای مدل داده‌ای برای این گروه‌ها عرضه شده است که می‌توان آن‌ها را در پنج گروه دسته‌بندی کرد (جدول 1).

بانگ‌های اطلاعاتی	رده‌بندی
Accumulo, Cassandra, Druid, HBase, Vertica	Column
Lotus Notes, Clusterpoint, Apache CouchDB, Couchbase, MarkLogic, MongoDB, OrientDB	Document
CouchDB, Oracle NoSQL Database, Dynamo, FoundationDB, MemcacheDB, Redis, Riak, c-treeACE, Aerospike, OrientDB, MUMPS	Key-value
Allegro, Neo4J, OrientDB, Virtuoso	Graph
OrientDB, FoundationDB, Alchemy Database	Multi-model

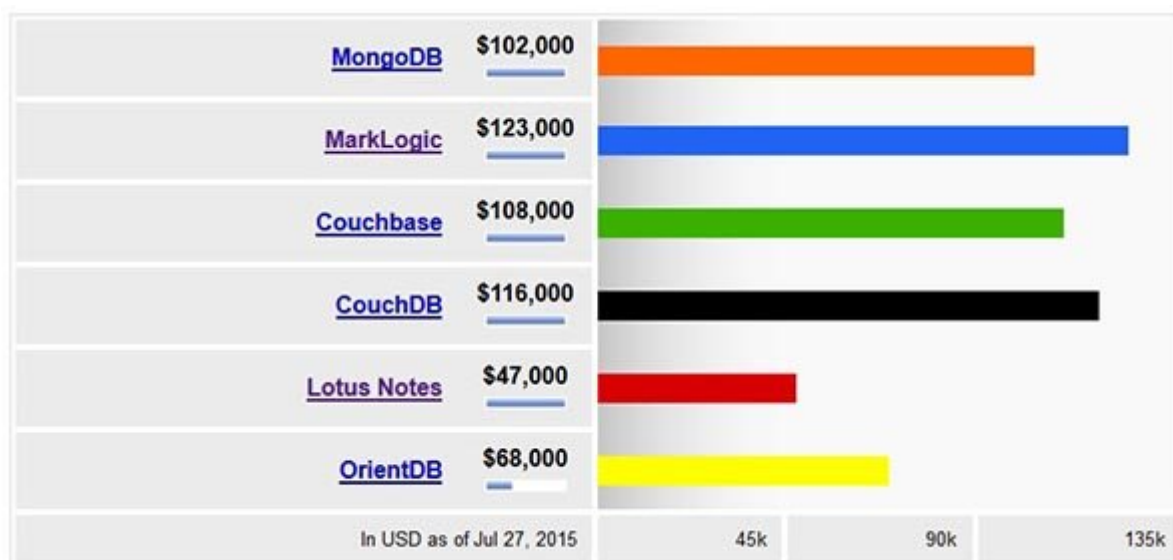
شکل 1: دسته‌بندی بانک‌های اطلاعاتی NoSQL

میزان دستمزد پرداختی برای هر یک از این گروه‌ها چگونه است؟
در گروه Column، بالاترین دستمزد پرداختی متعلق به HBase است (شکل 6).



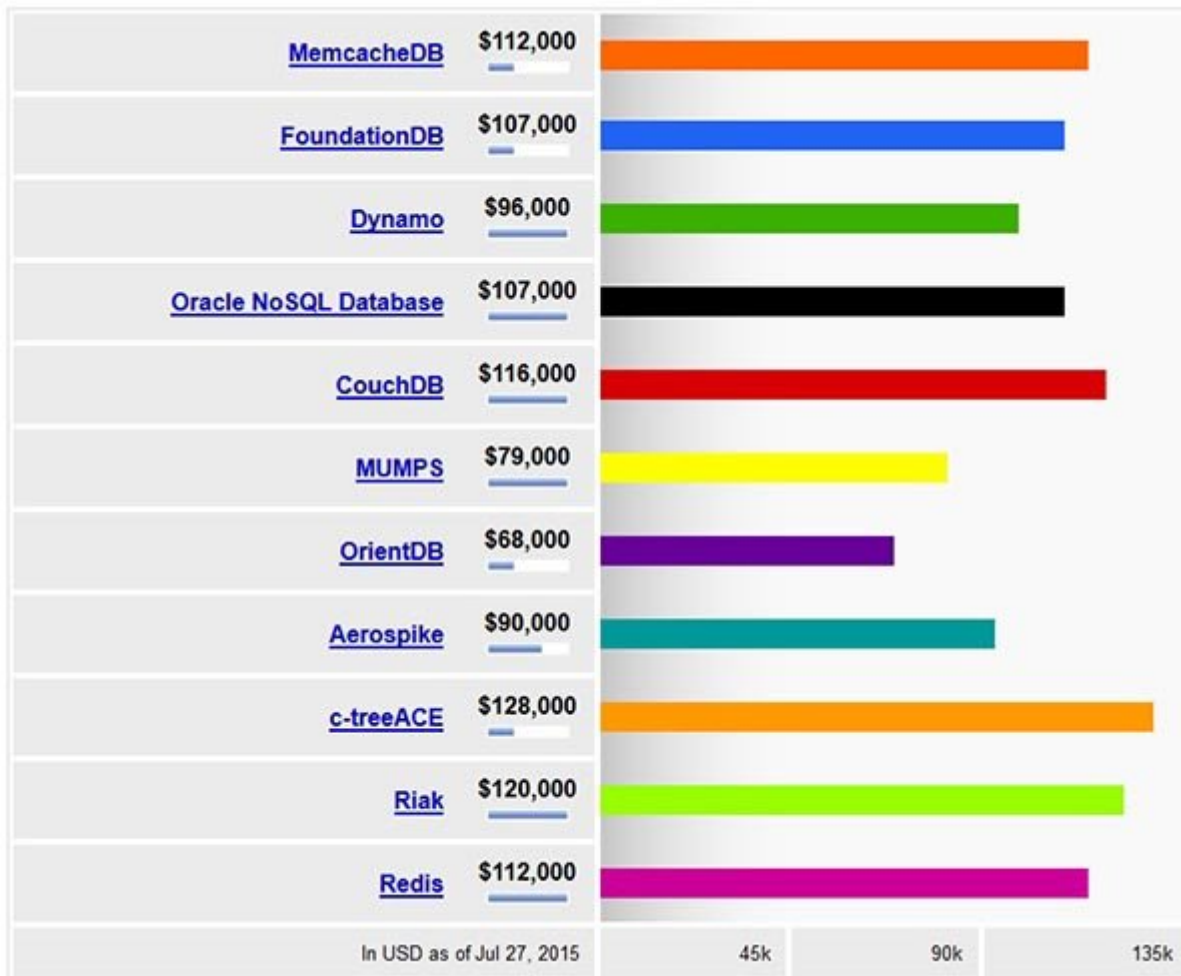
شکل 6: میانگین دستمزد پرداختی برای بانک‌های اطلاعاتی Column. میانگین دستمزد برای HBase و Druid به ترتیب 117 هزار دلار و 71 هزار دلار است (شکل 6).

در گروه Document، بالاترین دستمزد پرداختی متعلق به MarkLogic است (شکل 7).

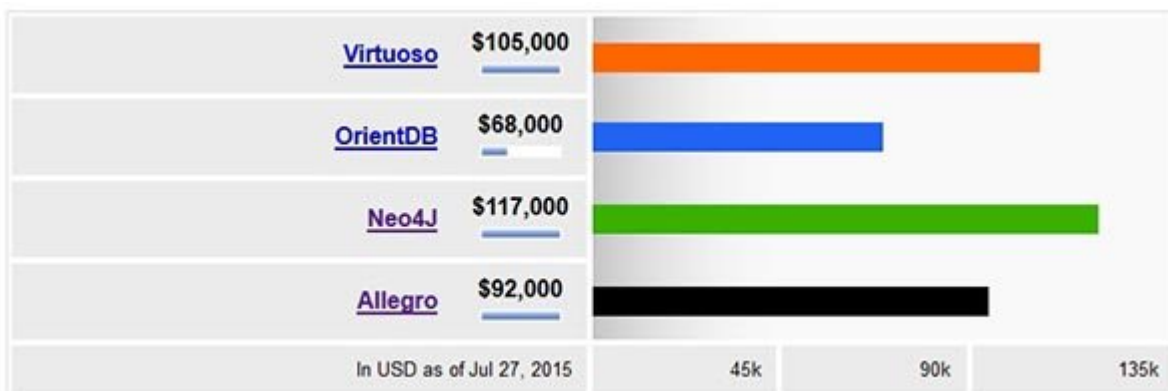


شکل 7: میانگین دستمزد پرداختی برای بانک‌های اطلاعاتی Document. میانگین دستمزد برای MarkLogic 123 هزار دلار است (شکل 7).

در گروه Key-value، بالاترین دستمزد پرداختی متعلق به C-treeACE است (شکل 8).



شماره 8: نمودار میله‌ای درج شده در جدول 86 نشان می‌دهد که *OrientDB* و *c-treeACE* در میان پایگاه‌های داده NoSQL با بالاترین درآمد دستمزد پرداختی متعلق به *Neo4J* است (شکل 9).



شماره 9: نمودار میله‌ای درج شده در جدول 71 نشان می‌دهد که *OrientDB* و *Neo4J* در میان پایگاه‌های داده Multi-model با بالاترین درآمد دستمزد پرداختی متعلق به *FoundationDB* است (شکل 10).



بن اسکا فیلد گروه بندی های مختلفی را برای بانک های اطلاعاتی در نظر گرفته است. کارایی بانک های اطلاعاتی را در جدول 2 مشاهده می کنید.

کارایی

بن اسکا فیلد گروه بندی های مختلفی را برای بانک های اطلاعاتی در نظر گرفته است. کارایی بانک های اطلاعاتی را در جدول 2 مشاهده می کنید.

مدل داده ای	عملکرد	مقیاس پذیری	انعطاف پذیری	پیچیدگی	قابلیت
Key-Value Store	یالا	یالا	یالا	تدارد	Variable (none)
Column-Oriented Store	یالا	یالا	متعادل	پایین	Minimal
Document-Oriented Store	یالا	متغیر (یالا)	یالا	پایین	Variable (low)
Graph Database	متغیر	متغیر	یالا	یالا	Graph theory
Relational Database	متغیر	متغیر	پایین	متعادل	Relational algebra

بن اسکا فیلد گروه بندی های مختلفی را برای بانک های اطلاعاتی در نظر گرفته است. کارایی بانک های اطلاعاتی را در جدول 2 مشاهده می کنید.

سخن آخر

در مجموع، شیوه کارکرد بانک های اطلاعاتی در آینده به طرز چشم گیری تغییر خواهد کرد. اگر تمایل دارید، کمی از وقت خود را صرف دنیای بانک های اطلاعاتی کنید. بهتر است به مفاهیمی همچون بزرگ داده، هوش تجاری و مفاهیم نوین دقت بیشتری داشته باشیم. بی شک شرکت ها در آینده به دنبال طراحانی خواهند بود که با این مفاهیم و ابزارهای مرتبط با آنها شگفتی ساز باشند.

تاریخ انتشار:
11 مهر 1394

نشانی منبع: <https://www.shabakeh-mag.com/cover-story/1783>