



دامستیک

انجمن علمی - دانشجویی گروه مهندسی علوم دامی دانشگاه تهران؛ بهار ۱۴۰۰

https://domesticsj.ut.ac.ir/article_82996.html

نرم افزار

معرفی قابلیت‌ها و کاربردهای نرم‌افزار آماری Statistica

شهگل رهبری^{۱*}

^۱ دانشجوی دکتری تخصصی تغذیه طیور، گروه مهندسی علوم دام و طیور، پردیس ابوریحان دانشگاه تهران، پاکدشت، ایران

چکیده

با توجه به پیچیدگی و حجم زیاد محاسبات آماری، استفاده از رایانه‌ها و نرم‌افزارهای محاسبات آماری، از ابزارهای اصلی متخصصان آمار و سایر افراد تحلیل‌گر داده‌های آماری محسوب می‌شوند. در این نوشتار به معرفی نرم‌افزار آماری Statistica پرداخته شده و خصوصیات و قابلیت‌های آن معرفی شده است. هر چند انواع مختلفی از نرم‌افزارهای محاسبات آماری دیگری مانند SAS یا SPSS نیز وجود دارند که به صورت بسته (package) نرم‌افزاری عمل می‌کنند؛ ولی شاید نوع نگاه Statistica به مفهوم و جنبه‌های دسته‌بندی تحلیل‌های آماری، عامل تمایز آن با سایر نرم‌افزارها است. نرم‌افزار Statistica یک بسته نرم‌افزاری تجزیه و تحلیلی آماری است. این نرم‌افزار آنالیز داده‌ها، مدیریت داده‌ها، داده‌کاوی و روش‌های تجسم داده‌ها را برای کاربر فراهم می‌کند. نرم‌افزار Statistica توسط شرکت StatSoft توسعه یافته است. قابلیت‌های کلیدی نرم‌افزار Statistica: تقسیم‌بندی مباحث علمی و نرم‌افزاری تحت عنوان ماژول‌های مربوط، قابلیت سفارشی‌شدن، خودکار عمل کردن و توسعه یافتن، قابلیت انجام آماره‌های توصیفی، ضریب همبستگی، محاسبه احتمالات، قابلیت رسم و ارائه بیش از صدها نوع نمایش گرافیکی با اتصال پویا بین نمودارها و داده‌ها (یعنی اتصالی که با تغییر داده‌ها نمودارها تغییر می‌کنند)، قابلیت برازش انواع توزیع‌ها به داده‌های عددی، قابلیت انجام آنالیز واریانس و کوواریانس یک طرفه و چند طرفه، انواع رگرسیون خصوصاً انجام رگرسیون با هزاران متغیر، قابلیت انجام آنالیز داده‌های ناپارامتری، قابلیت رسم نمودارهای کنترل کیفیت و تحلیل آن‌ها و نیز ویرایش این نمودارها، قابلیت تجزیه و تحلیل سری‌های زمانی و طراحی و تحلیل آزمایشات، قابلیت حل مسائل از طریق شبکه‌های عصبی، قابلیت یک سیستم خبره پیشنهاد کننده آسان و در دسترس، به طور کلی این نرم‌افزار به دلیل قابلیت‌های بالایی که در آنالیز داده دارد، می‌تواند در انجام تصمیم‌گیری و تحلیل داده‌ها به پژوهشگران علوم دامی کمک بسزایی نماید.

کلمات کلیدی: Statistica، آنالیز آماری، نرم‌افزارهای آماری، قابلیت نرم‌افزار، آنالیز داده

*نویسنده مسئول: shahgol.rahbari@gmail.com

تاریخ دریافت: ۱۴۰۰/۰۲/۰۴ تاریخ بازنگری: ۱۴۰۰/۰۲/۱۸ تاریخ پذیرش: ۱۴۰۰/۰۳/۰۱ تاریخ انتشار آنلاین: ۱۴۰۰/۰۳/۲۴



مقدمه

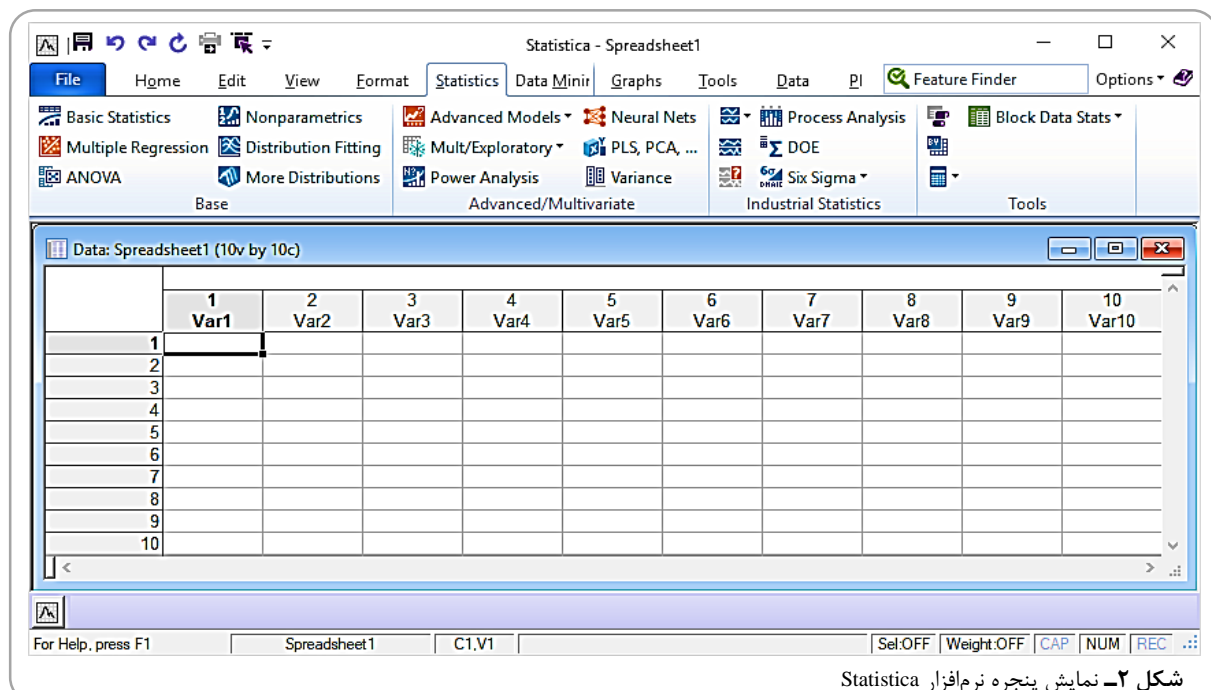
از ویژگی‌های مهم این نرم‌افزار می‌توان به موارد زیر اشاره

کرد:

- تحلیل داده‌ها
- مدیریت داده‌ها
- محاسبات آماری و آزمون‌های پارامتری و ناپارامتری چند متغیره
- محاسبات مربوط به کنترل کیفیت آماری
- داده‌کاوی
- یادگیری ماشین
- تحلیل متن
- تصویر سازی داده و رسم نمودارهای واضح و مؤثر

دستورات یا تحلیل‌ها در این نرم‌افزار از طریق انتخاب یکی از گزینه‌ها از فهرست دستورات امکان‌پذیر است. با ظاهر شدن پنجره مربوط به تحلیل مورد نظر، متغیرهای لازم برای تحلیل معرفی شده و نتیجه محاسبات طی جداول و نمودارهای مختلف آماری ظاهر می‌شود. این نتایج نیز قابل ذخیره‌سازی یا انتقال به نرم‌افزارهای دیگر مانند Word یا Excel هستند. در ادامه به فهرستی از توانایی‌های این نرم‌افزار اشاره خواهد شد.

این نرم‌افزار توسط شرکت StatSoft در سال ۱۹۸۰ توسعه یافت و براساس بسته‌های نرم‌افزاری مجزا یا بسته‌های محاسباتی مختلف به بازار عرضه شد. در سال ۱۹۸۶ نسخه کامل و به روز شده آن برای کامپیوترهای سیستم عامل شرکت اپل نیز آماده شده و نسخه سیستم عامل داس آن نیز که با نام امروزی Statistica به فروش می‌رفت، در سال ۱۹۹۱ تولید و به بازار آمد. ارائه گزارشات آماری در قالب جدول و نمودارهای آماری از خصوصیات بارز این نرم‌افزار در سال‌های ۱۹۹۰ تا به امروز بوده است. همچنین نسخه قابل اجرا در سیستم‌های ۳۲ و ۶۴ بیتی نیز در سال ۲۰۰۹ در دسترس کاربران قرار گرفت. این نسخه از Statistica به نام نسخه ۱۰ شهرت داشت. نسخه ۱۳،۰ این نرم‌افزار در سال ۲۰۱۴ ارائه شده و از آن به بعد این نرم‌افزار به تملک شرکت TIBCO درآمد و آن را به صورت مجتمع با نرم‌افزار Spotfire Data Science TIBCO در آورده است؛ ولی به صورت مجزا نیز قابل خریداری است.



شکل ۲- نمایش پنجره نرم‌افزار Statistica

طیف کاملی از تحلیل‌ها

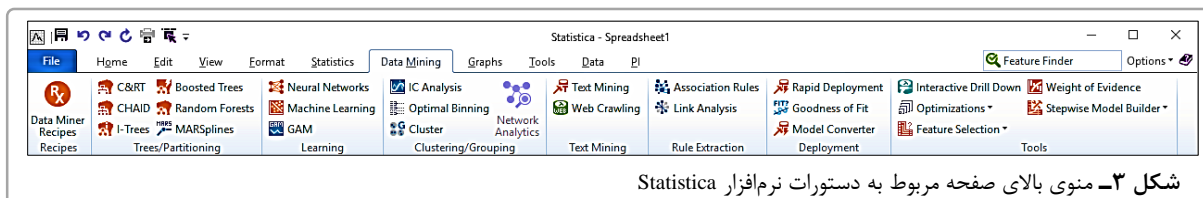
نرم‌افزار Statistica با بهره‌گیری از صدها تابع و الگوریتم‌های مختلف قادر است گزارشات آماری و محاسباتی را به راحتی انجام دهد که شامل موارد زیر هستند:

- طرح آزمایش‌ها
- تحلیل گراف/شبکه
- تحلیل متن

- تحلیل‌های سری زمانی
- رگرسیون
- خوشه‌بندی
- درخت تصمیم
- تحلیل آمار چندمتغیره
- کنترل فرآیند آماری

یادگیری ماشین روی کلان داده (Machine learning for big data)

براساس مدل‌های خودکاری که در Statistica قرار دارد، کاربر قادر خواهد بود که الگوریتم‌های مربوط به یادگیری ماشین را اجرا کرده و قابلیت‌ها و کارایی خود را در به کارگیری یادگیری ماشین روی کلان داده افزایش دهد. به این ترتیب می‌تواند به تشکیل الگوهای تصمیم‌گیری جدید بدون نیاز به نوشتن برنامه به زبان‌های R یا پایتون بپردازد. تصویر زیر دستورات مرتبط با یادگیری ماشین و داده‌کاوی در این نرم‌افزار را نشان می‌دهد.



شکل ۳- منوی بالای صفحه مربوط به دستورات نرم‌افزار Statistica

هماهنگی تیمی

انجام کار گروهی به کمک نرم‌افزار Statistica با منابع داده پویا، راحت‌تر صورت می‌گیرد. به این ترتیب چرخه تولید، تحلیل و به کارگیری الگوهای حاصل از داده‌ها ساده‌تر شده و گزارشات و نتایج حاصل معتبرتر و در زمان کوتاه‌تری آماده می‌شوند.

قابلیت رسم و ارائه بیش از صدها نوع نمایش گرافیکی با اتصال پویا بین نمودارها و داده‌ها (یعنی اتصالی که با تغییر داده‌ها نمودارها تغییر می‌کند).

قابلیت برآزش انواع توزیع‌ها به داده‌های عددی

قابلیت انجام آنالیز واریانس و کوواریانس یک‌طرفه و چند طرفه

انواع رگرسیون خصوصاً انجام رگرسیون با هزاران متغیر

قابلیت انجام آنالیز داده‌های ناپارامتری

قابلیت رسم نمودارهای کنترل کیفیت و تحلیل آن‌ها و نیز ویرایش این نمودارها

قابلیت تجزیه و تحلیل سری‌های زمانی و طراحی و تحلیل آزمایشات

قابلیت حل مسائل از طریق شبکه‌های عصبی

قابلیت یک سیستم خبره پیشنهاد کننده آسان و در دسترس

قابلیت گسترش و انعطاف‌پذیری

چرخه تولید داده و تحلیل‌های آماری با توجه به هماهنگی و امکان ارتباط Statistica با Jupyter Notebooks، پایتون، #C و همچنین R بسیار ساده است و تعامل با این موارد قدرت گسترش این نرم‌افزار را کامل می‌کند.

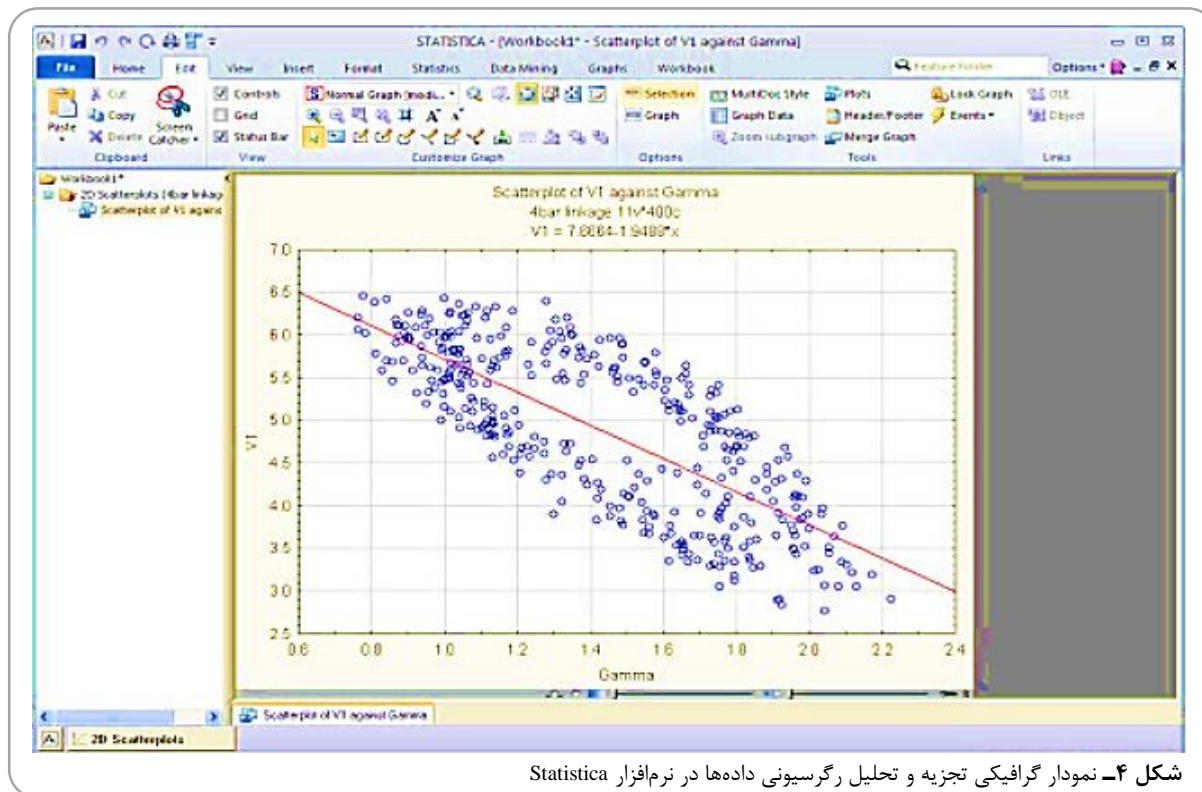
قابلیت‌های کلیدی نرم‌افزار Statistica

تقسیم‌بندی مباحث علمی و نرم‌افزاری تحت عنوان ماژول‌های مربوط

- قابلیت سفارشی‌شدن، خودکار عمل کردن و توسعه یافتن
- قابلیت انجام آماره‌های توصیفی، ضریب همبستگی، محاسبه احتمالات و ...

همزمان چند کلاینت بر روی یک سری داده و امکان انجام محاسبات تحت وب، با استفاده از اینترنت و یکپارچه سازی پایگاه‌های داده متفاوت در نقاط مختلف دنیا روی سرورهای متفاوت یک شرکت را فراهم می‌کند.

این نرم‌افزار به دلیل قابلیت‌هایی که دارد، مورد توجه بسیاری از شرکت‌های بزرگ صنعتی در دنیا قرار گرفت که نرم‌افزارهای دیگر از چنین توانایی‌هایی بی‌بهره هستند. مثلاً می‌توان از نسخه تحت شبکه این نرم‌افزار نام برد یا امکان آنالیز



شکل ۴- نمودار گرافیکی تجزیه و تحلیل رگرسیونی داده‌ها در نرم‌افزار Statistica

منابع

Marques de Sá, J. P. (2007). "Applied Statistics Using SPSS, STATISTICA, MATLAB and R". Spring, No. 2, New York, USA

نتیجه‌گیری کلی

به طور کلی می‌توان نتیجه گرفت که نرم‌افزار Statistica یک نرم‌افزار آماری است که می‌توان از آن برای آنالیز داده در همه رشته‌ها استفاده کرد. این نرم‌افزار به دلیل قابلیت‌های بالایی که در آنالیز داده دارد، مورد توجه پژوهشگران علوم دامی نیز بوده است. قابلیت رسم و ارائه بیش از صدها نوع نمایش گرافیکی با اتصال پویا بین نمودارها و داده‌ها (یعنی اتصالی که با تغییر داده‌ها نمودارها تغییر می‌کند) و انعطاف‌پذیری آن، سبب شده است تا کاربران در رشته علوم دامی تمایل به انجام آنالیزهای رگرسیونی مربوط به برآزش داده‌ها را در این نرم‌افزار داشته باشند. شناخت بیشتر این نرم‌افزار می‌تواند به کاربر در انجام آنالیزهایی مانند آنالیز واریانس و کوواریانس یک طرفه و چند طرفه، انواع رگرسیون خصوصاً انجام رگرسیون با هزاران متغیر در رشته علوم دامی کمک کند تا پژوهشگر بتواند به نتایج معتبرتری دست پیدا کند.

Publisher Note

Animal Science Students Scientific Association, Campus of Agriculture and Natural Resources at the University of Tehran

Submit Your Manuscript:

https://domesticstj.ut.ac.ir/contacts?_action=loginForm