

NOVINKY IBM SPSS STATISTICS

**Nové nástroje a funkce pro urychlení
a zjednodušení analýz**

Hlavní rysy:

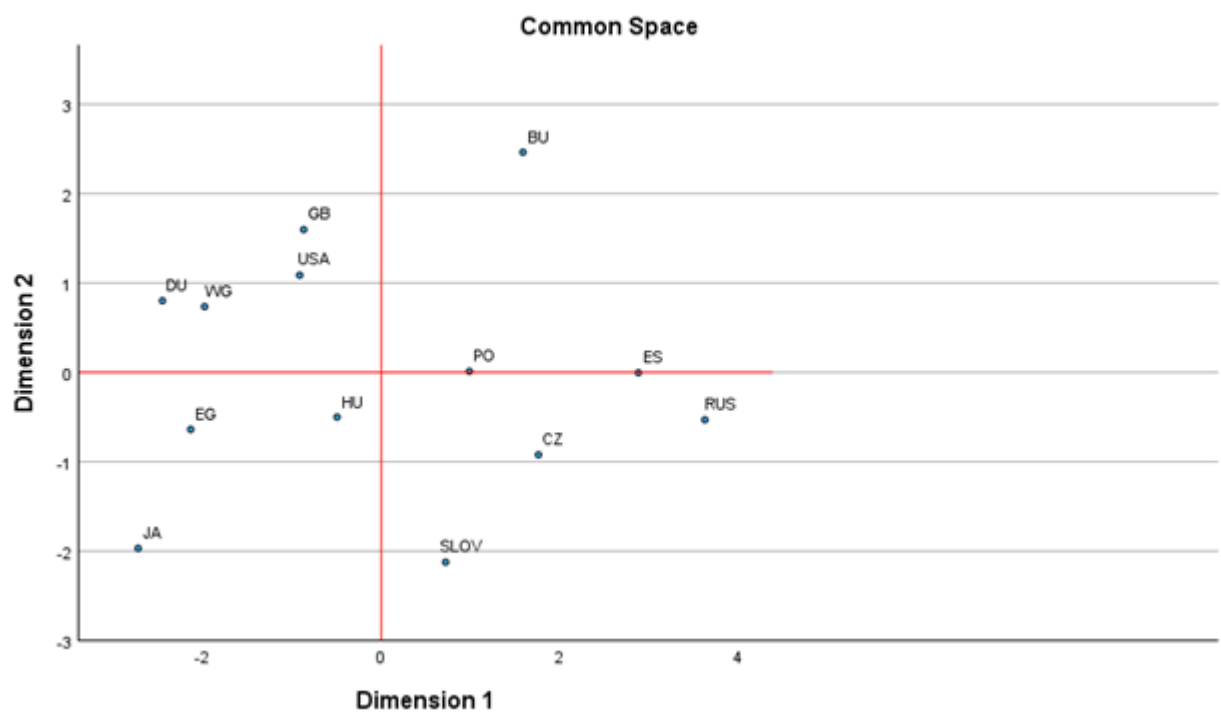
- Nové statistické procedury:
Proximity Mapping
Distance Correlation
Time Series Filters
Conditional Inference Tree
Multiple Adaptive Regression
Splines
Curated Help
- Zlepšení importu z Excelu
- Tvorba témat výstupů

Novinky IBM SPSS Statistics 31.0.0.0

Verze IBM SPSS Statistics 31 přidává dvě nové nativní procedury, tři procedury ve formě doplňků a několik vylepšení uživatelského rozhraní, což vede ke zlepšení uživatelského zážitku.

Proximity Mapping

Proximity Mapping je vizualizační technika, která prostorově organizuje vícerozměrná data a odhaluje vztahy mezi různými objekty. Tato procedura seskupuje podobné položky a zároveň odděluje rozdílné, čímž zpřehledňuje vzory v relačních datech. Procedura PROXMAP v IBM SPSS Statistics nabízí větší flexibilitu ve srovnání s PROXSCAL a dokáže zpracovat různé typy vstupů a vícero zdrojů údajů o blízkosti.



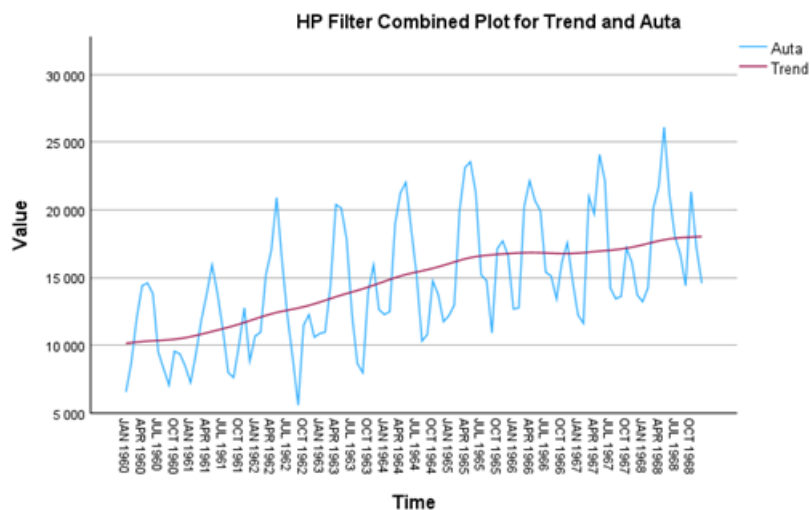
Distance Correlation

Distance Correlation detekuje lineární i nelineární závislosti. Nevyžaduje předpoklad normality a podporuje vícerozměrná data. Dokáže prokázat nezávislost proměnných přímočařeji než Pearsonova korelace.

Distance Correlation Estimates							
Variable Pair	dCov	dCor	Confidence Interval (Permutation, Sig. Level = 0.05)		95.00% Confidence Interval		Sig.
			Lower	Upper	Lower	Upper	
age vs. employ	0,0058	0,2799	0,0017	0,0071	0,2168	0,3408	<0,001
age vs. income	0,0018	0,2774	0,0019	0,0069	0,2141	0,3383	<0,001
age vs. debtinc	0,0000	0,0021	0,0017	0,0081	0,0000	0,0694	0,8600
employ vs. income	0,0026	0,4439	0,0018	0,0081	0,3883	0,4963	<0,001
employ vs. debtinc	0,0001	0,0052	0,0017	0,0092	0,0000	0,0725	0,1600
income vs. debtinc	0,0000	0,0036	0,0019	0,0080	0,0000	0,0709	0,3900

Time Series Filters

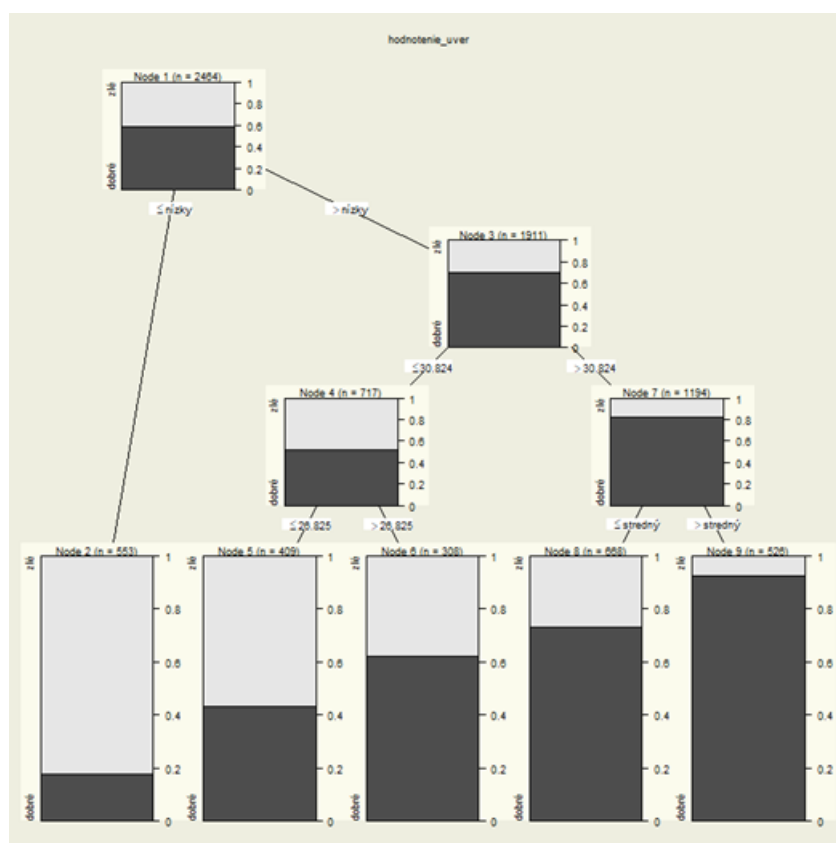
Procedura rozkládá časové řady na trendové a cyklické složky pomocí metod, jako jsou filtry Hodrick-Prescott, Baxter-King a Christiano-Fitzgerald. Tyto nástroje redukují šum, zvýrazňují klíčové vzory a podporují analýzu trendu, sezónnosti a reziduí. Procedura je dostupná jako doplněk a je nutné ji nainstalovat pomocí Extension Hub.



Source: Auta

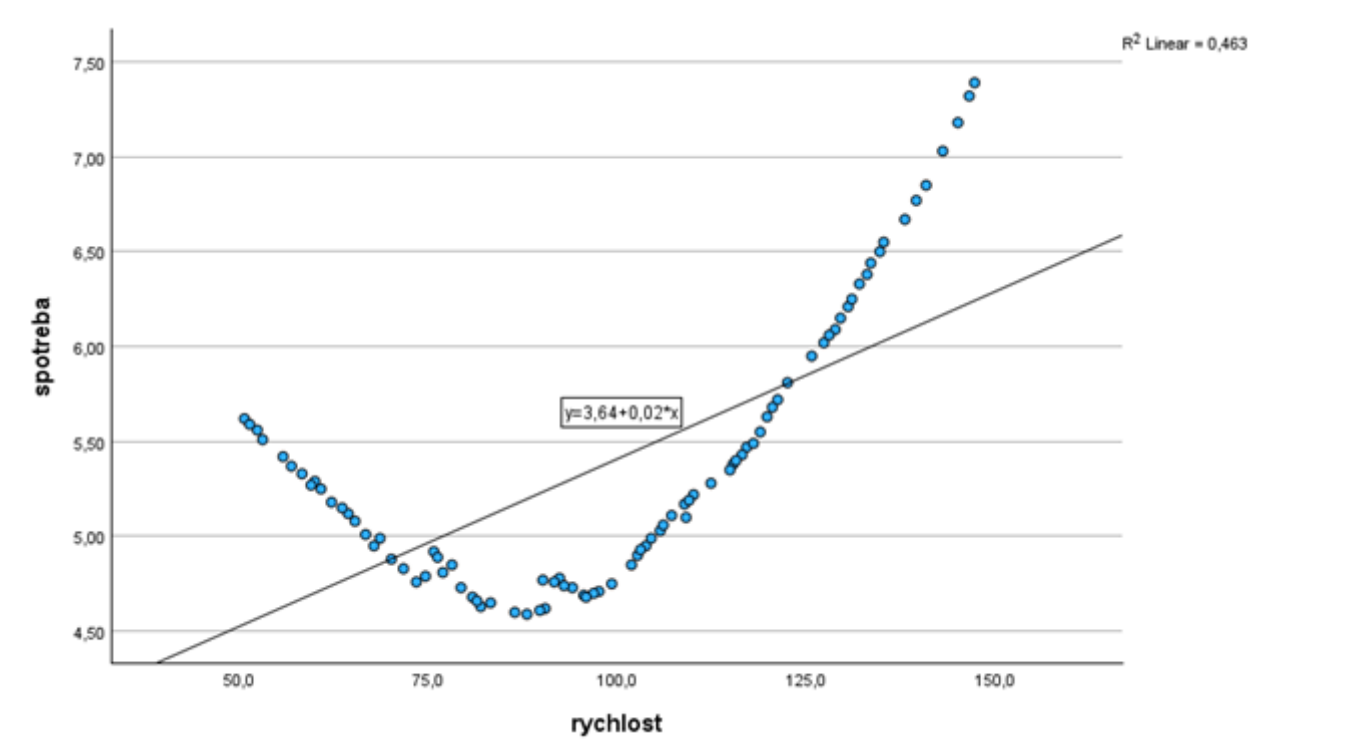
Conditional Inference Tree

Tyto stromy představují statisticky robustní alternativu k tradičním rozhodovacím stromům, čímž snižují pravděpodobnost přeučení a zlepšují interpretovatelnost díky použití statistických testů. Procedura poskytuje dva typy stromů – stromy podmíněné inference a stromy založené na modelech. Tato procedura je dostupná formou doplňku a je třeba ji nainstalovat přes Extension Hub.



Multiple Adaptive Regression Splines

Neparametrická regresní metoda, která automaticky identifikuje nejlepší místa pro body, kde model mění svou linearitu, což umožňuje lépe zachytit nelineární vztahy mezi proměnnými. Procedura je dostupná jako doplněk a vyžaduje instalaci pomocí Extension Hub.



Summary

Fit Summary					
	Generalized Cross Validation	Residual Sum of Squares	Standardized GCV	R Squared	Mean CV Rsq
1	0,003	0,211	0,994	0,995	0,989

Dependent variable: spotreba

Earth Coefficients

Coefficients	
	spotreba
(Intercept)	3,997
h(rychlost-95.6)	0,010
h(95.6-rychlost)	0,036
h(rychlost-122.5)	0,026
h(rychlost-73.5)	0,029

Dependent variable: spotreba
h() is the hinge function.
See dialog help

Curated Help

Analyzuje výstup pro vybrané procedury, které generují korelační koeficienty, a poskytuje souhrn klíčových zjištění barevným zvýrazněním tabulky a popisem jednotlivých barev. Toto vylepšení je dostupné pro procedury Bivariate Correlation, Partial Correlation, Distances, Canonical Correlation a Correlation in Linear Regression.

Correlations					
		Age in years	Years with current employer	Household income in thousands	Debt to income ratio (x100)
Age in years	Pearson Correlation	1	0,554**	0,476**	0,008
	Sig. (2-tailed)		<0,001	<0,001	0,810
	N	850	850	850	850
Years with current employer	Pearson Correlation	0,554**	1	0,625**	-0,034
	Sig. (2-tailed)	<0,001		<0,001	0,327
	N	850	850	850	850
Household income in thousands	Pearson Correlation	0,476**	0,625**	1	-0,036
	Sig. (2-tailed)	<0,001	<0,001		0,300
	N	850	850	850	850
Debt to income ratio (x100)	Pearson Correlation	0,008	-0,034	-0,036	1
	Sig. (2-tailed)	0,810	0,327	0,300	
	N	850	850	850	850

***. Correlation is significant at the 0.01 level (2-tailed).*

Pearson Correlations

- Highly Positive: (None)
- Positive: (Age in years <---> Years with current employer), (Age in years <---> Household income in thousands)
- No Linear Correlation: (None)
- Negative: (Years with current employer <---> Debt to income ratio (x100)), (Household income in thousands <---> Debt to income ratio (x100))
- Highly Negative: (None)

Note: Curated Help is calculated based on actual cell values, not the formatted values.

Další novinkou je přidaná procedura Chí-kvadrát testu, která je nyní dostupná přímo v nabídce Analýze. Vylepšen byl dialog procedury Independent Samples T Test, který nyní obsahuje ovládací prvek pro výběr testu homogenity rozptylu, a výstup v tomto případě obsahuje dvě tabulky.

Verze IBM SPSS Statistics 31 umožňuje přímo vypočítat variační koeficient ve vybraných procedurách, konkrétně Frequencies a Descriptives, a pomocí syntaxe i v procedurách ze skupin Complex Samples a Ratio.

Independent Samples Test									
t-test for Equality of Means									
		t	df	Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
				One-Sided p	Two-Sided p			Lower	Upper
Súčasný plat	Equal variances assumed	-10,945	472	<0,001	<0,001	-\$15,409.862	\$1,407.906	-\$18,176.401	-\$12,643.322
	Equal variances not assumed	-11,688	344,262	<0,001	<0,001	-\$15,409.862	\$1,318.400	-\$18,002.996	-\$12,816.728

Homogeneity of Variance Test		
Levene's Test for Equality of Variances		
	F	Sig.
Súčasný plat	119,669	<0,001

Statistics		
Súčasný plat		
N	Valid	474
	Missing	0
Mean		\$34,419.57
Median		\$28,875.00
Std. Deviation		\$17,075.661
Variance		291578214,45
CV		49,610

Další vylepšení ve verzi SPSS Statistics 31:

- Podpora obrázků na pozadí grafů.
- Zvýrazněná povinná pole v procedurách Power Analysis, Meta Analysis a Reports.
- Nová možnost zobrazit pouze hodnoty korelačních koeficientů pomocí skriptu dostupného v kořenovém adresáři Scripts - Reformat Correlations Table.py.



Vylepšený import z Excelu:

- Možnost výběru, zda se názvy proměnných načítají z prvního řádku souboru nebo z prvního řádku definovaného rozsahu.
- Hodnoty, které neodpovídají pravidlům pojmenování proměnných, se automaticky převedou na platné názvy a původní názvy se použijí jako popisky proměnných.

Read Excel File

C:\ACREA\Statistics\Novinky SPSS 31\Demo_data\Product.xlsx

Worksheet: Product [A1:B275]

Range:

Are variable names included at the top of your file :

☒ Yes

Row number that contains variable names: 1

☐ No

☒ Percentage of values that determine data type: 95

☒ Ignore hidden rows and columns

☐ Remove leading spaces from string values

☐ Remove trailing spaces from string values

Preview

Product	Product
99110	Aloe Relief
151110	Astro Pilot
151120	Astro Pilot
151130	Astro Pilot
153110	Auto Pilot
76110	Bear Edge
77110	Bear Survi...
122110	Bella

Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500 columns.

OK Paste Reset Cancel Help

Read Excel File

C:\ACREA\Statistics\Novinky SPSS 31\Demo_data\Product.xlsx

Worksheet: Product [A1:B275]

Range:

Are variable names included at the top of your file :

☒ Yes

Row number that contains variable names: 2

☐ No

☒ Percentage of values that determine data type: 95

☒ Ignore hidden rows and columns

☐ Remove leading spaces from string values

☐ Remove trailing spaces from string values

Preview

@99110	AloeR...
151110	Astro Pilot
151120	Astro Pilot
151130	Astro Pilot
153110	Auto Pilot
76110	Bear Edge
77110	Bear Survi...
122110	Bella
122120	Bella

Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500 columns.

OK Paste Reset Cancel Help

File Edit View Data Transform Analyze Graphs		
	@99110	AloeRelief
1	151110	Astro Pilot
2	151120	Astro Pilot
3	151130	Astro Pilot
4	153110	Auto Pilot

Dalším vylepšením je možnost tvorby témat výstupů, pomocí které si můžeme uložit preferované styly pro kontingenční tabulky a grafy do jednoho opakovaně použitelného tématu.

Options

General Language Viewer Data Currency **Output** Charts Pivot Tables File Locations Scripts Multiple Imputations Syntax Editor Privacy

Outline Labeling

Variables in item labels shown as:

Labels

Variable values in item labels shown as:

Labels

Pivot Table Labeling

Variables in labels shown as:

Labels

Variable values in labels shown as:

Labels

Curated Help

☒ Show Curated Help for Selected Pivot Tables

One Click Descriptives

☒ Suppress tables with many categories

Maximum number of categories: 24

☒ Include a chart in the output

Screen Reader Accessibility

☒ Read full row and column labels for each cell in pivot tables

☐ Read only row or column labels that change when you move to a different cell

Output Theme

Create Theme

Add

Default

MyTheme

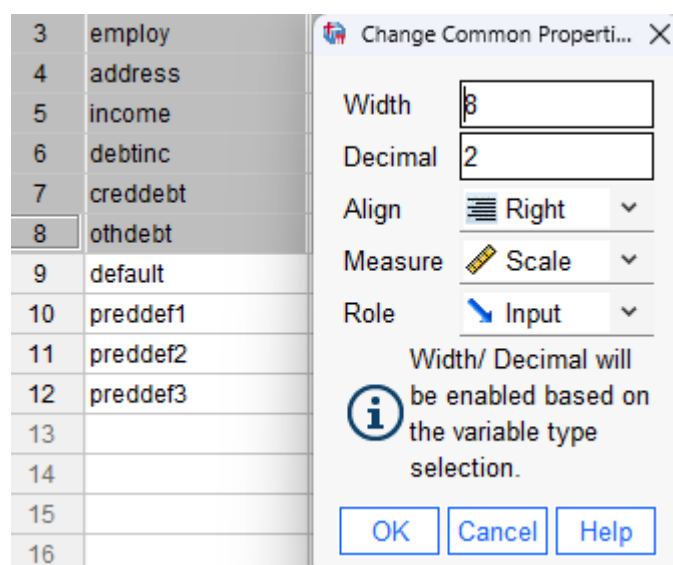
Change

Remove

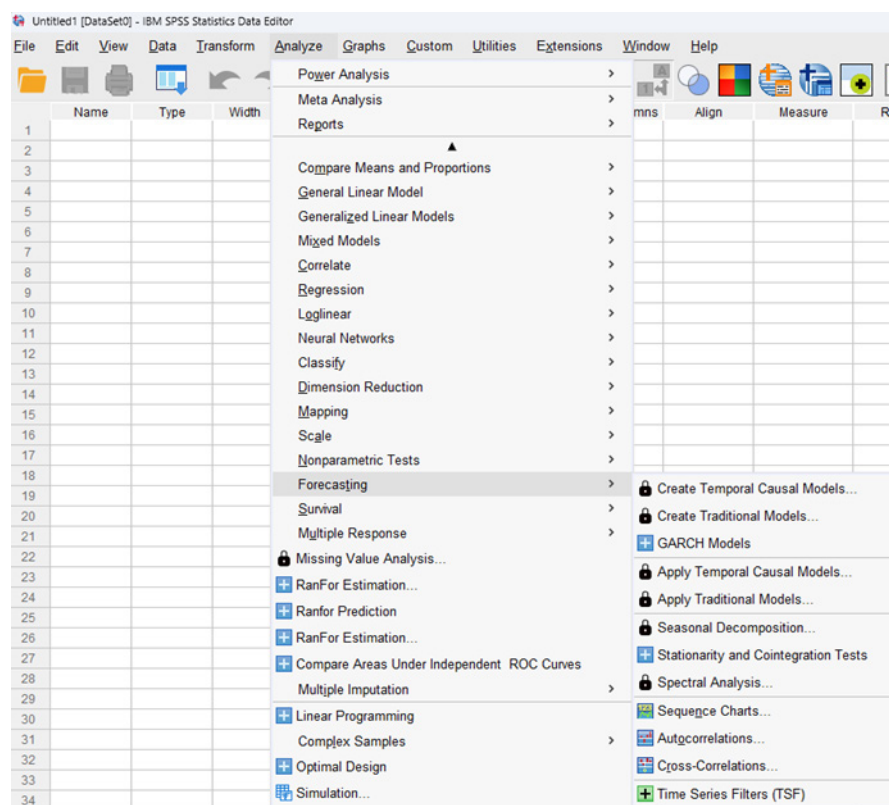
Group your preferred styles for Pivot Tables, Charts, and Viewer options into a single, reusable theme. Customize your settings in the Pivot Tables, Charts, and Viewer tabs. Enter a name for your new theme in the Create Theme field. Use Add to save your current settings as a new theme. Choose a theme from the list to apply its settings to your output. Select Change to modify a theme and Remove to delete it.

OK Cancel Apply Help

Přibyla možnost změny vlastností více atributů pro více proměnných současně a zobrazení počtu označených případů při aktivním filtru ve status bar.



Dalším vylepšením je zobrazení nelicencovaných procedur, které jsou označeny jako uzamčené ikony v nabídce.



Přidáno ve verzi 30.0.0.0

Bland-Altmanova analýza

Verze IBM SPSS Statistics 30 rozšiřuje své portfolio procedur o Bland-Altmanovu analýzu. Jedná se o neparametrickou metodu používanou k vyhodnocení shody mezi dvěma měřeními, ať už jde o porovnání nové metody s jinou metodou nebo s referenční normou.

Limity shody lze určit jako konkrétní percentil nebo jako dvě či tři standardní odchylky. Výsledkem je graf, který zobrazuje rozdíly mezi dvěma měřeními na ose Y oproti odpovídajícím hodnotám na ose X. Graf znázorňuje průměr dvou hodnot oproti jejich absolutnímu nebo relativnímu rozdílu. Tato procedura je součástí modulu Base.

Analýza normality

Verze 30 obsahuje možnost instalace nové procedury pro analýzu normality formou rozšíření, která je dostupná z Extensions Hub. Procedura rozšiřuje stávající univariační testy a grafy pro analýzu normality a přidává nové multivariační testy a grafy. Procedura umožňuje také identifikaci odlehklých hodnot pomocí kvantilové Mahalanobisovy vzdálenosti nebo upravené Mahalanobisovy vzdálenosti.

Multivariační testy normality

- Mardia test – robustní test založený na šikmosti a špičatosti
- Royston test – vhodný pro malé vzorky a nekorelované proměnné
- Henze-Zirkler test – založený na pozitivní vzdálenosti mezi teoretickou a empirickou distribuční funkcí
- Doornik-Hansen test – zohledňuje jak univariační, tak multivariační aspekty
- Energy test – robustní neparametrický test vhodný pro vysoce rozměrná data nebo při nesplnění parametrických předpokladů

Součástí multivariačních grafů jsou Chi-kvadrát kvantil-kvantil graf, perspektivní a obrysový graf.

Novinkou verze 30 je černé rozhraní, které uživatel nastaví v menu Edit – Options – General – Windows – Look and Feel.

Další novinkou je škálování textu na 4k monitorech, kde uživatel může zvětšit text od 100 % do 200 %. Škálování se aplikuje a menu, dialogová okna, datový editor a výstupní okno, konkrétně na tabulky, grafy a text.

Novinkou je také možnost zastavit procesor, pokud je zaneprázdněný, a v případě havárie je uživateli nabídnuta volba pro restart aplikace.

Vylepšen byl také panel OMS ve stavovém řádku, který je nyní aktivní a kliknutím na něj lze zastavit OMS proces. Panel filtr ve stavovém řádku umožňuje odstranit všechny aktivní filtry.

Nová verze obsahuje nové ikony aplikace a rychlejší spuštění díky nové verzi Javy.

Přidáno ve verzi 29.0.1

Významným vylepšením je přidání nové záložky Overview do datového okna, která poskytuje informace o datovém souboru. Jedná se o souhrn typů proměnných, podíly chybějících pozorování, základní analýza jednotlivých proměnných pomocí grafů a souhrnnými statistikami na základě definice úrovně měření.

Nová verze přináší parametrické regresní modely přežití pro údaje o opakujících se událostech zahrnutím pojmu sdílené křehkosti. S tímto pojmem se zachází jako s náhodnou složkou, aby se zohlednil nepozorovaný účinek v důsledku individuální nebo skupinové variability.

Přibyla samostatná procedura pro výpočet percentilů, včetně specifikace kvartilů nebo vlastních percentilů s možností výběru z pěti metod odhadu.

Součástí novinek je i vylepšení více procedur:

- **Linear Regression**
 - » PRESS (Predicted residual sum of squares) – nabízí možnost výpočtu populární statistice pro hodnocení modelů lineární regrese
 - » Selection Criteria – ukazatel kvality regresního modelu
 - » Tolerance – definování úrovně pro zpracování proměnných vykazujících téměř kolinearitu
- **ROC Curves** - Youden's Index, který kombinuje senzitivitu a specifitu do jediné míry diskriminace v každém bodě řezu na křivce ROC
- **Linear a Generalized Linear Mixed Models** přibyla možnost exportu predikcí náhodných efektů do nových datových souborů – EBLUPs
- **Explore** – výpočet percentilů se stejnou funkcionalitou jako samostatná procedura Percentiles

- **Cox w/ Time-Dep Cov** – možnost specifikace/výpočtu více časově závislých kovariátů

Samozřejmostí je vylepšení syntaxe výše uvedených procedur.

Součástí novinek jsou i praktická zlepšení, která se týkají úpravy výstupu - Output Modify. Pravým tlačítkem myši na výstupní tabulku získáme rychlý přístup k některým běžným funkcím jako je transpozice, řazení podle sloupce, skrytí sloupce a zvýraznění buněk ve sloupci.

Vylepšeno bylo i dialogové okno Search, které nově prohledává i proměnné, komunitu IBM SPSS a také Youtube kanál IBM SPSS.

Novinkou je také dokumentace v českém jazyce. Jedná se konkrétně o instalační manuály jak pro klientské, tak pro serverové verze a dokumentaci k modulům.

Přidáno ve verzi 29

- Nové statistické procedury: Modely zrychleného času selhání, Alternativy lineární MNČ.
- Vylepšení procedur Lineárních smíšených modelů a v Zobecněných lineárních smíšených modelech.
- Nový typ grafu Violin Plot.
- Zlepšení praktické použitelnosti.

Modely zrychleného času selhání (Accelerated Failure Time, AFT)

Verze IBM SPSS Statistics 29 přináší nový model do procedur analýzy přežití. Na rozdíl od stávajících tabulek přežití, Kaplan-Meierovy a Coxovy regrese, má nově přidáný model zrychleného času selhání parametrický charakter. To znamená, že se předpokládá, že závislá proměnná má specifické rozdělení. Parametrické modely se často považují za méně flexibilní než neparametrické modely, ale pokud závislá proměnná má identifikovatelné rozdělení, mohou být tyto druhy postupů velmi účinné. Zatímco modely proporcionálních rizik předpokládají, že účinkem kovariátu je znásobení rizika nějakou konstantou, model AFT předpokládá, že účinky kovariátu zkracují nebo prodlužují dobu přežití o nějakou konstantu. Tato schopnost může být užitečná pro výzkumníky, kteří zkoumají čas do selhání jako součást režimu preventivní údržby, zejména pokud je známo, že faktory jako míra využití zařízení urychlují nebo zpomalují čas do selhání.

Nová procedura podporuje parametrické modely založené na Weibullově, exponenciálním, log-normálním a log-logistickém rozdělení. Procedura je součástí modulu Regression.

Alternativy lineární MNČ: Lasso, Ridge a Elastic Net (Linear OLS Alternatives)

Verze 29 obsahuje také možnost instalace tří nových regresních procedur, které využívají různé formy regularizace. Procedury Lasso, Ridge a Elastic Net mohou být přidány do podnabídky Regression. Pomocí knihovny strojového učení Python Scikit-learn, lze tyto postupy rychle nainstalovat jako rozšíření.

Všechny tyto techniky jsou optimalizovány, aby se předešlo problémům s multikolinearitou, které jsou nezdědky spojeny s metodou nejmenších čtverců. Obecně tyto regulační techniky fungují tak, že penalizují velké modelové koeficienty.

Lasso

Procedura Lasso (Least Absolute Shrinkage and Selection Operator) často označována jako regularizace L1, funguje tak, že penalizuje nejméně důležité parametry a zmenšuje je k nule. Užitečná je při výběru proměnných, protože slabé proměnné se efektivně vynulují, čímž se zjednoduší konečný model.

Hřebenová regrese

Regularizace L2, známá jako Hřebenová regrese, má tendenci penalizovat koeficienty rovnoměrnějším způsobem než L1. Kromě vytváření jednodušších modelů se běžně používá při řešení otázek multikolinearity.

Elastic Net

Elastic Net kombinuje Lasso (L1) a hřebenovou regresi (L2), což může vést k vyváženějšímu modelu, je-li každá jednotlivá metoda nějakým způsobem suboptimální.

Míry Pseudo-R² v Lineárních smíšených modelech a v Zobecněných lineárních smíšených modelech

Výstup z lineárních smíšených modelů a zobecněných lineárních smíšených modelů nyní zahrnuje míry pseudo-R² a vnitroskupinový korelační koeficient - ICC. R² je běžně uváděná statistika označující podíl rozptylu vysvětleného lineárním modelem. Vnitroskupinový korelační koeficient (ICC) je statistika, která uvádí, jak velká část reziduálního rozptylu se vysvětluje skupinovým faktorem ve víceúrovňových/hierarchických modelech.

Podpora vyšších verzí Python a R

S verzí IBM SPSS Statistics 29 je nainstalován Python 3.10.4 a R 4.2.0

Odfiltrované případy lze znovu zobrazit

Odfiltrované případy již nejsou skryty v editoru dat, když je vybrána podmnožina případů a nevybrané případy se nesmažou. Jedná se o návrat chování SPSS Statistics 27.0.1 a starších verzí.

Violin Plot

Violin Plot byly přidány do nabídky Grafboard Template Chooser. Tyto grafy jsou kombinací krabicových grafů a grafů hustoty jádra. Violin Plot zobrazují vrcholy dat a používají se pro vizualizaci rozdělení spojitých proměnných. Na rozdíl od krabicového grafu, který může zobrazovat pouze souhrnné statistiky, Violin Plot zobrazují souhrnné statistiky a hustotu každé proměnné.

Přidáno ve verzi 28

- Nová statistická procedura meta-analýza
- Vylepšení procedur Power analýza a Ratio statistics
- Nový typ grafu Relationship map
- Zlepšení každodenní použitelnosti

Meta-analýza (Meta analysis)

Novinkou verze 28 je nová statistická procedura Meta-analýza (meta analysis). Meta-analýza je vědecká metoda, která souhrnně analyzuje data z několika na sobě nezávislých studií. Cílem je identifikace a kvantifikace převažujících trendů nebo zjištění příčin rozdílných závěrů studií. IBM SPSS Statistics 28 představuje celkem 5 nových procedur. Mimo jiné obsahují testy heterogenity, meta-regresní analýzu, či několik nových typů grafů pro lepší vizualizaci výsledků.

Seznam procedur:

- Meta-analysis of continuous outcomes: Raw data, precalculated effect size data
- Meta-analysis of binary outcomes: Raw data, precalculated effect size data
- Meta-regression

Power analýza

Power analýza, která byla představena ve verzi 27 přichází s novým vylepšením procedury. Power analýza vám dovolí určit optimální rozsah výběru pro vaši studii nebo projekt. Pokud je získání dat drahé, je důležité vědět, kolik pozorování je třeba získat, aby se dosáhlo dostatečné pravděpodobnosti

detekce určeného efektu. Novinkou verze 28 je možnost zadat několik samostatných hodnot požadované síly (power). Hodnoty mohou být zadány buď jednotlivě nebo je možné zadat celý rozsah, ve kterém se mají pohybovat. Tato novinka byla přidána do všech procedur Power analýzy.

Ratio

Nejzásadnější změnou procedury v nové verzi je přidání nové statistiky PRB (price-related bias) a příslušný interval spolehlivosti. V hlavním okně procedury je možnost zahrnout počet případů, ze kterých byly příslušné statistiky počítány do výstupové tabulky.

Relationship map

Jednou z hlavních novinek ve verzi 28 je nový typ grafu, tzv. Relationship map. Pomocí tohoto typu grafu lze vizuálně zkoumat vzájemné vztahy mezi několika kategorizovanými proměnnými. Graf vizuálně znázorňuje vzájemné vazby pomocí uzlů a spojnic. Uzly představují proměnné a jejich kategorie, tloušťka spojnic mezi uzly reprezentuje sílu vztahu mezi nimi. Graf je možné najít na záložce Graphs -> Relationship Map.

Další vylepšení procedur

V uživatelském rozhraní procedury GLM (General Linear Model) nyní navíc přibyl pod volbou EM Means nové zaškrtnávací políčko Compare simple main effects. Procedura One-way ANOVA nyní podporuje nečíselné kategorizované proměnné.

Zlepšení každodenní použitelnosti

Verze 28 představuje nové funkce a rozšíření, jejichž cílem je zlepšit každodenní použitelnost a celkovou produktivitu.

Workbook

IBM SPSS Statistics 28 přichází s novým typem okna Workbook. Mimo jiné umožňuje toto okno zároveň spuštění syntaxe a zobrazení jejího výsledku. Další výhodou je možnost plné editace syntaxe, výstupů nebo i textu, přímo v tomto okně. Soubory tohoto formátu mají příponu .spwb. Pro aktivaci je nutné změnit nastavení pod volbou Edit > Options > Application Mode na Workbook

Side-panel table editor

Verze 28 představuje nové vylepšení editace tabulek. Editační okno nyní obsahuje boční panel, který zahrnuje všechny

funkce z nabídky Cell Properties a Table Properties, které byly původně obsaženy pod volbou Format, což značně zkrátí a zjednoduší čas strávený editací tabulek.

Search

Funkce Search, která byla poprvé představena ve verzi 27, nyní prohledává všechna slova / výrazy v každém dialogovém okně uživatelského rozhraní.

High contrast

Do SPSS Statistics 28 byla přidána podpora pro režim Vysokého kontrastu. Pokud zvolíte režim Vysoký kontrast a otevřete SPSS Statistics, bude se automaticky zobrazovat v tomto režimu. Chcete-li použít vysoký kontrast pro počítače s Mac OS, je nutné vybrat položku Systém Preferences > Accessibility a tažením posuvníku Kontrastu displeje upravíte kontrast.

Přidáno ve verzi 27

- nové procedury: Power Analysis,
- změny v modulu Base: příprava dat a bootstrapové funkce,
- vylepšení procedur: Effect size, Weighted Kappa a MATRIX a kvantilová regrese.

Změny v modulu Base

Od verze IBM SPSS Statistics 27 jsou do modulu Base zahrnuty moduly Data Preparation a Bootstrapping. Daná změna poskytuje všem zákazníkům Statistics 27 přístup k pokročilým funkcím z těchto dvou modulů.

Power analysis

Power analýza hraje klíčovou roli při plánování experimentů. SPSS Statistics 27 nově nabízí 11 procedur pro power analýzu, které Vám pomohou určit optimální rozsah výběru pro Vaši studii. Procedury umí spočítat nejen potřebný rozsah výběru pro požadovanou sílu, ale i sílu testu pro daný rozsah. K dispozici jsou i grafické možnosti, např. zobrazení síly jako funkce tzv. effect size nebo 3-dimenzionální grafy zobrazující sílu jako funkci effect size a rozsahu výběru.

Weighted Kappa

Nově zařazená statistická metoda Weighted Kappa se používá pro vyjádření shody mezi dvěma hodnotiteli, kteří hodnotí na ordinální škále, a umožňuje použít váhy pro různou míru neshody.

Effect size

Tzv. effect sizes zobrazují míru porušení nulové hypotézy. O výpočet effect size jsou nově obohacené dialogy t-testů a ANOVA.

MATRIX a kvantilová regrese

IBM SPSS Statistics 27 přináší také vylepšení procedury MATRIX o nové funkce a rozšířila se nabídka procedury kvantilové regrese o možnost efektivněji zadávat mřížku kvantilů.

Další změny

- nová funkce Search v nabídce Edit,
- bublinový graf jako čtyřdimenzionální scatterplot,
- kompatibilita programu s vyšší verzí Pythonu 3.8.2.

Přidáno ve verzi 26

- nové Bayesovské metody,
- vylepšená integrace s produkty třetích stran (Microsoft Graphic Object),
- možnost úpravy grafů již při jejich tvorbě,
- přidání klávesových zkratk do syntax editoru.

Bayesovská statistika

Bayesovská statistika je založena na odlišném principu než klasická (frekventistická) statistika a může sloužit jako alternativa, nebo další možnost k analýze našich dat. Některé procedury využívající bayesovský přístup v IBM SPSS Statistics jmenovitě jsou:

- T-test (jednovýběrový, párový, dvouvýběrový),
- Pearsonův korelační koeficient,
- lineární regrese,
- analýza rozptylu jednoduchého třídění.

Vylepšení grafiky

Díky nové možnosti exportu, Microsoft Graphic Object, se mnohem lépe manipuluje s grafy v programech MS Office. Tyto grafy se pak chovají jako by byly vytvořeny přímo v prostředí MS Office a jejich následná úprava je výrazně jednodušší.

Dalším vylepšením prošel tzv. Chart Builder, kde přibyly další možnosti editace grafu již při jeho tvorbě a celkově byla vylepšena vizualizace grafických výstupů. Jmenovitě můžeme vybírat barvy kategorií pro kategorizované proměnné, slučovat méně četné kategorie do jedné nebo upravovat samotnou velikost grafu.

Syntax editor

Změny v syntax editoru jsou mířeny hlavně na uživatele, kteří s ním běžně pracují. Zápis syntaxe se dočkal úpravy ve formě klávesových zkratk, které zjednodušují psaní, formátování i editaci. Pomocí klávesových zkratk lze řádky rychleji spojovat, duplikovat, odstraňovat, přesunovat atp. Asi nejvýznamnější úpravou je možnost kopírování dat z datového editoru či MS Excelu do více řádků najednou. Tyto novinky jsou dostupné v modulu Base.

Další změny

Některé další úpravy jsou například:

- předělání dialogového okna pro spojování souborů,
- zjednodušení panelu nástrojů,
- možnost otevírat a ukládat soubory Stata v14.

Přidáno ve verzi 25

Bayesovská statistika je založena na odlišném principu než klasická (frekvencistická) statistika. Může sloužit jako další možnost k analýze našich dat. Některé procedury využívající bayesovský přístup v IBM SPSS Statistics jmenovitě jsou:

- T-test (jednovýběrový, párový, dvouvýběrový),
- Pearsonův korelační koeficient,
- lineární regrese,
- analýza rozptylu jednoduchého třídění.

Díky nové možnosti exportu, Microsoft Graphic Object, se mnohem lépe pracuje s grafy v programech MS Office. Tyto grafy se pak chovají jako by byly vytvořeny přímo v prostředí MS Office a jejich následná úprava je výrazně jednodušší.

Dalším vylepšením prošel tzv. Chart Builder, kde přibyly další možnosti editace grafu již při jeho tvorbě a celkově byla vylepšena vizualizace grafických výstupů. Jmenovitě můžeme vybírat barvy kategorií pro kategorizované proměnné, slučovat méně četné kategorie do jedné nebo upravovat samotnou velikost grafu.

Změny v syntaxovém editoru jsou mířeny hlavně na uživatele, kteří s ním běžně pracují. Zápis syntaxe se dočkal úpravy ve formě klávesových zkratk, které zjednodušují psaní, formátování i editaci. Pomocí klávesových zkratk lze řádky rychleji spojovat, duplikovat, odstraňovat, přesunovat atp. Významnou úpravou je možnost kopírování dat z datového editoru či MS Excelu do více řádků najednou. Tyto novinky jsou dostupné v modulu Base.

Některé další úpravy jsou například:

- předělání dialogového okna pro spojování souborů,
- zjednodušení panelu nástrojů,
- možnost otevírat a ukládat soubory Stata v14.

Přidáno ve verzi 24

- rozšiřující doplňky z webu,
- snadnější tvorba a sdílení vlastních doplňků,
- nový průvodce načítáním dat,
- vylepšení v modulu IBM SPSS Custom Tables.

V IBM SPSS Statistics 24 naleznete nový způsob práce se zdrojem programových doplňků od třetích stran na **Statistics Extension Hub** a nástroj **Custom Dialog Builder for Extensions** pro snadnou práci s doplňky.

The Extensions Hub je nové prostředí pro správu a získání doplňků, kde najdete, stáhnete, aktualizujete, případně odinstalujete doplněk založený na syntaxi IBM SPSS, která může přistupovat k programu Python nebo R. Prostřednictvím IBM SPSS Statistics Custom Dialog Builder for Extensions je správa, vytváření a sdílení doplňků založených na R a Python velmi jednoduchá - snadno a rychle vytvoříte doplněk v grafickém rozhraní pro tvorbu dialogových oken, doplněk můžete sdílet, instalovat a umístit do nabídek programu.

Způsob **načítání a zápisu dat** z externích zdrojů ve verzi 24 obsahuje lepší rozpoznávání typu proměnných, vynechání skrytých sloupců a řádků v MS Excel, odstranění mezer ze začátků a konců textových řetězců, export popisků hodnot do databáze a mnoho dalšího.

Modul IBM SPSS Custom Tables obsahuje několik vylepšení. Součástí je např. designová váha (effective weights) pro výběry s nestejnou pravděpodobností, zobrazení intervalů spolehlivosti a směrodatných odchylek statistik, zobrazení výsledků signifikantních testů přímo v těle tabulky a zobrazení signifikancí pro testy sloupcových průměrů a proporcí.

Přidáno ve verzi 23

Nové algoritmy

- modelování vztahů mezi velkým počtem časových řad pomocí časového kauzálního modelování (Temporal Casual Modeling)

Geoprostorové analýzy

- modely vztahů napříč časem a prostorem pomocí časoprostorové předpovědi (Spatial Temporal Prediction)
- zobecněná prostorová asociační pravidla zahrnující geoprostorové atributy pro předpověď výskytu událostí (Generalized Spatial Association Rule)

Zvýšená efektivita

- vylepšená kategorická analýza hlavních komponent obsahující neparametrický bootstrapping a nové rotace pro vylepšení konvergence
- vylepšené možnosti vytváření vlastních skriptů pomocí jazyka R ve vývojovém prostředí zabudovaném v SPSS pro snadnější vytváření a ladění skriptů
- nová generace webových výstupů s vyšší interaktivitou, lepší funkcí a podporou webových serverů
- načítání a ukládání souborů Stata 9-13
- lepší možnosti importovat, číst a zapisovat soubory Stata 9-13
- rychlejší zápis do databáze na základě dávkového nahrávání (bulk export/insert)
- podpora přístupu k softwaru pomocí identifikačních karet

Přidáno ve verzi 22

- interaktivní zobrazení výstupu na chytrých zařízeních
- příprava výstupů v prezentační kvalitě
- efektivnější modelování pomocí Monte Carlo simulací, podpora automatického lineárního modelování (Automatic Linear Modeling), simulace textových proměnných a teplotních grafů
- rychlejší nalezení informací na úvodní stránce
- zvýšení výkonu a škálovatelnosti pomocí IBM SPSS Server a SQL pushback
- instalace Python plugin v rámci hlavní instalace SPSS Statistics

- vyhledávání, stahování a instalace dostupných skriptů přímo z SPSS Statistics., možnost vyhledávání aktualizace již nainstalovaných skriptů a stahování a instalování jejich aktualizované verze
- zjednodušená metoda specifikace uživatelských odhadů v SPSS Amos
- zdokonalená podpora přihlašování pro podnikové normy (Enterprise Standard) v normách platformy (Platform Standards)
- možnost číst/zapisovat šifrované datové soubory SPSS Statistics prostřednictvím jiných aplikací pomocí vstupně-výstupní dli knihovny
- výstup neparametrických testů formou standardních pivotních tabulek

Přidáno ve verzi 21

- Monte Carlo simulace pro vytváření lepších modelů pro nejisté vstupy
- porovnání dvou datových souborů nebo datových matic pro zajištění kompatibility hodnot a informací o proměnných
- zabezpečení citlivých datových a výstupních souborů před prohlížením nekompetentními osobami pomocí šifrování a hesla
- vylepšené a rychlejší spojování dat
- zdokonalení pivotních tabulek novými funkcemi jako je např. snadná navigace a třídění
- jednodušší specifikace modelu v IBM SPSS Amos, nástroji pro strukturální modelování vztahů (SEM)
- programování v SPSS Statistics pomocí Java pluginu
- možnost importu dat z IBM® Cognos® Business Intelligence do SPSS Statistics pro podrobnější analýzy
- export výstupů do Microsoft Excel 2007/2010
- lepší škálovatelnost a vyšší výkon pomocí „load balancing“ v IBM SPSS Collaboration and Deployment Services
- systém jednotného přihlášení (single sign-on) mezi klientem SPSS Statistics a serverem
- pro prostředí s omezeným zabezpečením: možnost spouštět SPSS Statistics Server jako „non-root“ uživatel na operačních systémech UNIX/LINUX
- komprese datových souborů pro uchování místa na disku
- kompatibilita novějších verzí klienta se staršími verzemi serveru a naopak (např. klient verze 21 a server verze 20 nebo klient verze 20 a server verze 21)

Přidáno ve verzi 20

- mapové šablony a podpora ESRI souborů v IBM SPSS Statistics Base
- rychlejší tvorba a práce s pivotními tabulkami
- GLMM procedura v IBM SPSS Advanced Statistics i pro ordinální závislé proměnné
- specifikace modelů v IBM SPSS Amos formou jednoduchého programového kódu
- spouštění dávek IBM SPSS Statistics Server v offline modu bez nutnosti připojení klienta na server
- komprese dočasných souborů vznikajících při třídění pomocí IBM SPSS Statistics Server pro ušetření místa na disku zejména při třídění velkých souborů